STIC-Biotech/ChemLib

From: Chan, Christina
Sent: Wednesday, October 01, 2003 4:35 PM
To: Schnizer, Holly; STIC-Biotech/ChemLib
Subject: RE: Request for RUSH sequence search for Appl. no. 09/444,281

Importance: High

Please rush. Thanks Chris

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
308-3973
CM-1, 9B19

-----Original Message-----
Fr m: Schnizer, Holly
Sent: Wednesday, October 01, 2003 4:26 PM
To: Chan, Christina
Subject: Request for RUSH sequence search for Appl. no. 09/444,281

I would like to request the following RUSH sequence search for the above appl. which is an amended due this biweek (Oct. 6).

Please search the commercial and interference databases for the peptide of SEQ ID NO:85 (ILPWKWPWWPWRR)

Thank you.

Holly Schnizer
AU 1653
CM1-9E09
305-3722
mailbox: CM1-9B01

Searcher:_____
Phone:_____
Location:_____
Date Picked Up:_____
Date Completed:_____
Searcher Prep/Review:_____
Clerical:_____
Online time:_____

TYPE OF SEARCH:
NA Sequences:_____
AA Sequences:_____
Structures:_____
Bibliographic:_____
Litigation:_____
Full text:_____
Patent Family:_____
Other:_____

VENDOR/COST (where applic.)
STN:_____
DIALOG:_____
Questel/Orbit:_____
DRLink:_____
Lexis/Nexis:_____
Sequence Sys.:_____
WWW/Internet:_____
Other (specify):_____

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 2, 2003, 09:56:42 ; Search time 77 Seconds
(without alignments)
26.798 Million cell updates/sec

Title: US-09-444-281-85
Perfect score: 99
Sequence: 1 ILPWKWPWWPWRR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	99	100.0	13	14	AAR30970	Broad spectrum ant
2	99	100.0	13	16	AAR78457	Indolicidin analog
3	99	100.0	13	19	AAY24608	Indolicidin analog
4	99	100.0	13	19	AAW66441	Cationic peptide i
5	99	100.0	13	20	AAW87609	Antimicrobial pept
6	99	100.0	13	21	AAY92794	Synthetic antimicr
7	99	100.0	13	21	AAY91740	Cationic peptide I
8	99	100.0	13	21	AAY91771	Amino acid sequenc
9	99	100.0	13	21	AAY91772	Amino acid sequenc
10	99	100.0	13	21	AAY44666	Crosslink-stabilis
11	99	100.0	13	21	AAY44324	Antimicrobial pept
12	99	100.0	13	21	AAY55056	Non-amidated indol
13	99	100.0	13	21	AAY57123	Naturally occurrin
14	99	100.0	13	22	ABP60382	Indolicidin peptid
15	99	100.0	13	22	ABP60383	Indolicidin peptid
16	99	100.0	13	22	AAB91842	Antimicrobial pept
17	99	100.0	13	23	ABB81940	Peptide fragment o
18	99	100.0	13	23	ABP59052	Peptide #1. Synth
19	99	100.0	13	23	AAO15561	L-indolicidin carr
20	99	100.0	13	23	ABB81249	Indolicidin antiba
21	99	100.0	13	23	ABB81261	Indolicidin antiba
22	99	100.0	13	23	ABB07699	Bovine cathelicidi
23	99	100.0	13	23	AAU90977	Transplant media a
24	99	100.0	13	24	ABG76068	Human regulatory p
25	99	100.0	13	24	AAE34433	Cow indolicidin pe
26	99	100.0	13	24	ABU59617	Cationic cancer -t
27	99	100.0	13	24	ABR00800	Bioactive syntheti
28	99	100.0	13	24	ABR00815	Bioactive syntheti
29	99	100.0	14	21	AAY57118	Indolicidin peptid
30	99	100.0	14	21	AAY57143	Indolicidin peptid
31	99	100.0	15	18	AAW12879	Antimicrobial cati
32	99	100.0	15	24	ABG73946	Cell wall/cell mem
33	99	100.0	16	21	AAY57144	Indolicidin peptid
34	99	100.0	19	23	AAB47907	C-terminus of ubiq
35	99	100.0	63	21	AAY44668	Poly-(Indol (1-13)
36	99	100.0	63	21	AAY57142	Indolicidin fusion
37	99	100.0	144	23	ABB07706	Bovine peptide ant
38	96	97.0	13	16	AAR78459	Indolicidin analog
39	96	97.0	13	22	ABP60388	Indolicidin peptid
40	96	97.0	15	18	AAW12878	Antimicrobial cati
41	95	96.0	12	21	AAY44669	Amino terminal tru
42	95	96.0	12	21	AAY57110	Indolicidin peptid
43	95	96.0	12	22	ABP60391	Indolicidin peptid
44	94	94.9	12	16	AAR78458	Indolicidin analog
45	94	94.9	12	21	AAY57133	Indolicidin peptid

ALIGNMENTS

RESULT 1
AAR30970
ID AAR30970 standard; peptide; 13 AA.
XX
AC AAR30970;
XX
DT 25-MAR-2003 (updated)
DT 12-MAY-1993 (first entry)
XX
DE Broad spectrum antimicrobial indolicidin peptide.
XX
KW Tryptophan rich; microbial; microbistatic; inhibition.
XX
OS Bos taurus.
XX
PN WO9222308-A1.
XX
PD 23-DEC-1992.
XX
PF 10-JUN-1992; 92WO-US04920.
XX
PR 14-JUN-1991; 91US-0715271.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Cullor JS, Selsted ME;
XX
DR WPI; 1993-017896/02.
XX
PT Broad spectrum antimicrobial cpd. obtd. from bovine granulocytes -
PT comprises tryptophan rich peptide, pref. having low immunogenicity
PT and comprising proline rich peptide or carboxy terminal amide

XX
PS Claim 2; Page 19; 29pp; English.
XX
CC The sequence is that of an indolicidin peptide which shows broad
CC spectrum antimicrobial activity and when administered to a host does
CC not elicit an immune response. It is effective against viruses, gram
CC positive bacteria, gram negative bacteria and fungi, including
CC Staphylococcus aureus, Escherichia coli, Salmonella typhimurium,
CC Listeria monocytogenes, Candida albicans and Cryptococcus neoformans.
CC It can be used as a therapeutic agent, food preservative or
CC disinfectant, e.g. to purify a water supply. The peptide is pref.
CC administered at an effective amt. of 0.5-500 ug/ml final concentration.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 99; DB 14; Length 13;
Best Local Similarity 100.0%; Pred. No. 9.8e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKWPWWPWRR 13
||||||||||||
Db 1 ILPWKWPWWPWRR 13

RESULT 2
AAR78457
ID AAR78457 standard; peptide; 13 AA.
XX
AC AAR78457;
XX
DT 25-MAR-1996 (first entry)
XX
DE Indolicidin analog #4.
XX
KW Indolicidin; microbicide; therapeutic agent; prophylactic;
KW food preservative; disinfectant; medication; Gram positive bacteria;
KW Gram negative bacteria; protozoa; yeast; fungi; viruses.
XX
OS Synthetic.
XX
PN WO9522338-A1.
XX
PD 24-AUG-1995.
XX
PF 10-FEB-1995; 95WO-US01895.
XX
PR 16-FEB-1994; 94US-0197205.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Selsted ME;
XX
DR WPI; 1995-302552/39.
XX
PT Analogues of the tryptophan-rich peptide indolicidin - exhibiting
PT broad spectrum antimicrobial activity and selectivity without
PT undesirable side effects
XX
PS Claim 6; Page 27; 37pp; English.
XX
CC The sequences represented by AAR78454-R78459 are indolicidin analogues.
CC These analogues exhibit broad spectrum antimicrobial activity and have
CC antimicrobial selectivity when compared to naturally occuring
CC indolicidin. The antimicrobial activity of these analogues can be
CC altered by incorporation of D-form, chemically altered or synthetic
CC amino acids. These sequences can be incorporated into a pharmaceutical
CC composition (e.g. as a liposome or non-liposome lipid complex carrier)
CC for use in a microbicidal method. These sequences are active against
CC Gram positive and negative bacteria, protozoa, yeast, fungi and viruses.
CC They can be used as therapeutic agents, prophylactics, food
CC preservatives, disinfectants or medications. These sequences are easily
CC synthesised in an active and effective broad spectrum antimicrobial form
CC with decreased undesirable side effects. Compared to naturally occuring
CC indolicidin, these analogues show increased antimicrobial and decreased
CC haemolytic activity. Peptide stability, and period of activity within
CC the cell can be increased or decreased according to the incorporation of
CC D- or L-form amino acids.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 99; DB 16; Length 13;
Best Local Similarity 100.0%; Pred. No. 9.8e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKWPWWPWRR 13
||||||||||||
Db 1 ILPWKWPWWPWRR 13

RESULT 3
AAY24608
ID AAY24608 standard; peptide; 13 AA.
XX
AC AAY24608;
XX
DT 18-AUG-1999 (first entry)
XX
DE Indolicidin analogue #60.
XX
KW Indolicidin; bacterial infection; photo-oxidised solubiliser;
KW antimicrobial; antibiotic; antiarrythmic; surface disinfectant;
KW additive; shampoo; soap; insecticide; herbicide; preservative;
KW food; technical material.
XX
OS Synthetic.
XX
PN WO9807745-A2.
XX
PD 26-FEB-1998.
XX
PF 21-AUG-1997; 97WO-US14779.
XX
PR 13-JAN-1997; 97US-0034949.
PR 21-AUG-1996; 96US-0024754.
XX
PA (MICR-) MICROLOGIX BIOTECH INC.
XX
PI Erfle D, Fraser JR, Krieger TJ, Taylor R, West MH;
XX
DR WPI; 1998-169090/15.
XX
PT New indolicidin analogues with antimicrobial activity and related
PT nucleic acid - vectors, transformed cells and antibodies, also
PT conjugates with polyoxyalkylene glycol and fatty acid to reduce
PT toxicity, useful therapeutically, as disinfectants etc.
XX
PS Example 1; Page 32; 129pp; English.
XX
CC AAY24549 to AAY24615 represent indolicidin analogues of formulae
CC (I)-(VIII) containing up to 25 amino acids (aa): RXZXXZXB (I), BXZXXZXB
CC (II), BBBXZXXZXB (III), BXZXXZXBBBn(AA)nMILBBAGS (IV), BXZXXZXBB(AA)nM
CC (V), LBBnXZnXXZnXRK (VI), LKnXZXXZXRRK (VII) and BBXZXXZXBBB (VIII).
CC Where Z = P or V; X = hydrophobic residue, preferably W; B = basic aa,
CC preferably R or K; AA = any aa; n = 0 or 1; in (II), at least 1 Z = V;
CC in (VIII) at least 2 X = F or Y. The analogues are used to treat
CC infections caused by bacteria (Gram positive or negative, or anaerobic);
CC fungi (yeast or moulds); parasites (protozoa, nematodes, cestodes or
CC trematodes) or viruses. Typical of very many pathogens that can be
CC controlled are Leishmania, Trypanosoma, Ascaris lumbricoides, Fasciola
CC hepatica, Klebsiella pneumoniae, Bordetella pertussis, Staphylococcus
CC aureus, Listeria, Clostridium, rotavirus and papilloma virus. Compounds
CC derived from the analogues may be used similarly; the compounds may
CC also be prepared from antibiotics or antiarrythmic agents. The analogues
CC may be used therapeutically or to coat medical devices; also they are
CC useful as surface disinfectants, as additives to shampoo or soaps, as

CC insecticides or herbicides, or as preservatives for foods and technical
CC materials. The analogues are administered by injection, lavage, orally
CC or topically, generally at 0.1-50 mg/kg. These analogues have a broader
CC spectrum of activity than indolicidin and modification as compounds
CC reduces their toxicity.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 99; DB 19; Length 13;
Best Local Similarity 100.0%; Pred. No. 9.8e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKWPWWPWRR 13
||||||||||||||
Db 1 ILPWKWPWWPWRR 13

RESULT 4
AAW66441
ID AAW66441 standard; peptide; 13 AA.
XX
AC AAW66441;
XX
DT 12-JAN-1999 (first entry)
XX
DE Cationic peptide indolicidin.
XX
KW Indolicidin analogue; resistance; cationic peptide; antibiotic;
KW bacterial infection; tolerance; antibacterial; microorganism;
KW bacteria; fungus; parasite; virus.
XX
OS Bos taurus.
XX
PN WO9840401-A2.
XX
PD 17-SEP-1998.
XX
PF 10-MAR-1998; 98WO-CA00190.
XX
PR 25-FEB-1998; 98US-0030619.
PR 10-MAR-1997; 97US-0040649.
PR 20-AUG-1997; 97US-0915314.
PR 26-SEP-1997; 97US-0060099.
XX
PA (MICR-) MICROLOGIX BIOTECH INC.
XX
PI Fraser JR, McNicol PJ, West MHP;
XX
DR WPI; 1998-520800/44.
XX
PT New indolicidin peptide analogues - useful for, e.g. enhancing
PT activity of antibiotic or overcoming tolerance, acquired resistance
PT or inherent resistance of microorganisms
XX
PS Disclosure; Page 10; 105pp; English.
XX
CC AAW66393 to AAW66469 represent native cationic peptides from the
CC present invention. The present invention describes compositions and
CC methods for treating infection, especially bacterial infections. The
CC compositions and methods use cationic peptides in combination with an
CC antibiotic agent which are then administered to a patient to enhance the
CC activity of the antibiotic agent, to overcome: (a) tolerance; (b)
CC acquired resistance; and (c) inherent resistance. The combinations of
CC antibiotics and cationic peptides can provide synergistic activity
CC against a microorganism that is tolerant, inherently resistant, or has
CC acquired resistance to an antibiotic agent. They can be used for killing
CC e.g. bacteria, fungi, parasites and viruses.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 99; DB 19; Length 13;
Best Local Similarity 100.0%; Pred. No. 9.8e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKWPWWPWRR 13
||||||||||||||
Db 1 ILPWKWPWWPWRR 13

RESULT 5
AAW87609
ID AAW87609 standard; peptide; 13 AA.
XX
AC AAW87609;
XX
DT 19-MAR-1999 (first entry)
XX
DE Antimicrobial peptide Indolicidin.
XX
KW Antimicrobial; fusion; acidic peptide; recombinant; microorganism;
KW guamerin; basic peptide; Indolicidin.
XX
OS Bos sp.
XX
PN WO9854336-A1.
XX
PD 03-DEC-1998.
XX
PF 28-MAY-1998; 98WO-KR00132.
XX
PR 09-APR-1998; 98KR-0013372.
PR 28-MAY-1997; 97KR-0021312.
XX
PA (KOAD) KOREA ADV INST SCI & TECHNOLOGY.
PA (SAMY-) SAMYANG GENEX CORP.
XX
PI Hong S, Kang MH, Kim JH, Kim S, Lee H, Lee JH;
XX
DR WPI; 1999-059844/05.
DR N-PSDB; AAV83788.
XX
PT New method for mass production of antimicrobial peptides - by
PT constructing fusion genes comprising acidic and antimicrobial
PT peptide genes and transforming host with vector containing these
XX
PS Example 6; Page 18; 52pp; English.
XX
CC The invention relates to mass production of antimicrobial peptides. The
CC method comprises constructing a fusion gene containing a first gene
CC encoding a negatively charged acidic peptide having at least two cysteine
CC residues, and a second gene encoding a positively charged basic
CC antimicrobial peptide. A host microorganism is transformed with a vector
CC containing the fusion gene and then cultured. The expressed antimicrobial
CC peptide is then recovered. The method is used to mass produce
CC antimicrobial peptides in recombinant microorganisms. The inhibitory
CC effect of the expressed antimicrobial peptide upon the growth of the host
CC microorganism is considerably reduced by fusing it to the acidic peptide.
CC Therefore, the use of the fusion gene provides an economic, recombinant
CC alternative of mass producing antimicrobial peptides, which overcomes the
CC disadvantages of low-productivity and poor economy, previously
CC encountered by recombinant and chemical methods. The present sequence
CC represents an antimicrobial peptide Indolicidin. The encoding DNA
CC can be used along with the acidic peptide Guamerin gene in the
CC construction of the fusion gene.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 99; DB 20; Length 13;
Best Local Similarity 100.0%; Pred. No. 9.8e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKWPWWPWRR 13
||||||||||||||
Db 1 ILPWKWPWWPWRR 13

```
RESULT 6
AAY92794
ID   AAY92794 standard; peptide; 13 AA.
XX
AC   AAY92794;
XX
DT   29-AUG-2000  (first entry)
XX
DE   Synthetic antimicrobial peptide, indolicidin.
XX
KW   Magainin; antimicrobial; transgenic plant; protease degradation; Rev4;
KW   indolicidin; protein production; reverse peptide.
XX
OS   Bos taurus.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   13
FT                   /note= "amidated"
XX
PN   WO200026344-A1.
XX
PD   11-MAY-2000.
XX
PF   29-OCT-1999;   99WO-US25561.
XX
PR   30-OCT-1998;   98US-0106373.
PR   02-NOV-1998;   98US-0106537.
XX
PA   (INTE-) INTERLINK BIOTECHNOLOGIES LLC.
PA   (KENT ) UNIV KENTUCKY RES FOUND.
XX
PI   Everett NP,  Li Q,  Lawrence C,  Davies MH;
XX
DR   WPI; 2000-365597/31.
XX
PT   Polypeptides for reducing proteolytic degradation of proteins
PT   administered to, or produced by a plant comprise indolicin or its
PT   functional equivalents
XX
PS   Example 2; Page 15; 50pp; English.
XX
CC   Indolicidin is a potent antimicrobial tridecapeptide, originally
CC   purified from cytoplasmic granules of bovine neutrophils. A reverse
CC   peptide, Rev4 (AAY92796) of indolicidin was found to have increased
CC   stability against plant protease degradation. Expression of antimicrobial
CC   peptides in transgenic plants suffers a major limitation in that the
CC   foreign peptides are susceptible to rapid degradation by proteases. The
CC   invention concerns reducing the extent of protease degradation of a
CC   protein applied to, or produced by a plant by administering indolicidin,
CC   Rev4 or a functional equivalent to the plant. Transgenic plants
CC   expressing indolicidin and Rev4 are useful for production of the
CC   antimicrobial peptides. Compositions containing indolicidin and Rev4 are
CC   also useful for production of agronomically important proteins in
CC   plants.
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 99;  DB 21;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 9.8e-07;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPWRR 13
              |||||||||||||
Db          1 ILPWKWPWWPWRR 13


RESULT 7
AAY91740
ID   AAY91740 standard; Peptide; 13 AA.
XX
AC   AAY91740;
XX
DT   06-JUN-2000  (first entry)
XX
DE   Cationic peptide Indolicidin amino acid sequence.
XX
KW   Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
KW   leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;
KW   breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
KW   multidrug resistance.
XX
OS   Unidentified.
XX
PN   WO9965506-A2.
XX
PD   23-DEC-1999.
XX
PF   14-JUN-1999;   99WO-CA00552.
XX
PR   12-JUN-1998;   98US-0096541.
XX
PA   (MICR-) MICROLOGIX BIOTECH INC.
XX
PI   Friedland HD,  Krieger TJ,  Taylor R,  Erfle D,  Fraser JR,  West MHP;
XX
DR   WPI; 2000-223549/19.
XX
PT   Novel pharmaceutical composition containing optionally activated
PT   polyoxyalkylene-modified cationic peptides, useful for treating tumours
PT   -
XX
PS   Disclosure; Page 11; 94pp; English.
XX
CC   This sequence represents a cationic peptide amino acid sequence, which
CC   can be used in the pharmaceutical composition of the invention. The
CC   invention relates to a pharmaceutical composition containing at least one
CC   activated polyoxyalkylene (APO)-modified cationic peptide. The
CC   modification of peptides with APO increases their activity against tumour
CC   cells, including those with a multidrug resistant phenotype. The
CC   pharmaceutical composition can be used to treat tumours, specifically
CC   lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
CC   cervix, uterus, skin, prostate, liver and colon.
XX
SQ   Sequence   13 AA;

  Query Match             100.0%;  Score 99;  DB 21;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 9.8e-07;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPWRR 13
              |||||||||||||
Db          1 ILPWKWPWWPWRR 13


RESULT 8
AAY91771
ID   AAY91771 standard; Peptide; 13 AA.
XX
AC   AAY91771;
XX
DT   06-JUN-2000  (first entry)
XX
DE   Amino acid sequence of cationic peptide MBI 10.
XX
KW   Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
KW   leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;
KW   breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
KW   multidrug resistance.
XX
OS   Synthetic.
XX
PN   WO9965506-A2.
XX
PD   23-DEC-1999.
XX
```

PF 14-JUN-1999; 99WO-CA00552.
XX
PR 12-JUN-1998; 98US-0096541.
XX
PA (MICR-) MICROLOGIX BIOTECH INC.
XX
PI Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;
XX
DR WPI; 2000-223549/19.
XX
PT Novel pharmaceutical composition containing optionally activated
PT polyoxyalkylene-modified cationic peptides, useful for treating tumours
PT -
XX
PS Disclosure; Page 14; 94pp; English.
XX
CC This sequence represents a cationic peptide amino acid sequence, which
CC can be used in the pharmaceutical composition of the invention. The
CC invention relates to a pharmaceutical composition containing at least one
CC activated polyoxyalkylene (APO)-modified cationic peptide. The
CC modification of peptides with APO increases their activity against tumour
CC cells, including those with a multidrug resistant phenotype. The
CC pharmaceutical composition can be used to treat tumours, specifically
CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
CC cervix, uterus, skin, prostate, liver and colon.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 99; DB 21; Length 13;
Best Local Similarity 100.0%; Pred. No. 9.8e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKWPWWPWRR 13
||||||||||||||
Db 1 ILPWKWPWWPWRR 13

RESULT 9
AAY91772
ID AAY91772 standard; Peptide; 13 AA.
XX
AC AAY91772;
XX
DT 06-JUN-2000 (first entry)
XX
DE Amino acid sequence of cationic peptide MBI 10CN.
XX
KW Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
KW leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;
KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
KW multidrug resistance.
XX
OS Synthetic.
XX
PN WO9965506-A2.
XX
PD 23-DEC-1999.
XX
PF 14-JUN-1999; 99WO-CA00552.
XX
PR 12-JUN-1998; 98US-0096541.
XX
PA (MICR-) MICROLOGIX BIOTECH INC.
XX
PI Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;
XX
DR WPI; 2000-223549/19.
XX
PT Novel pharmaceutical composition containing optionally activated
PT polyoxyalkylene-modified cationic peptides, useful for treating tumours
PT -
XX
PS Disclosure; Page 14; 94pp; English.
XX
CC This sequence represents a cationic peptide amino acid sequence, which
CC can be used in the pharmaceutical composition of the invention. The
CC invention relates to a pharmaceutical composition containing at least one
CC activated polyoxyalkylene (APO)-modified cationic peptide. The
CC modification of peptides with APO increases their activity against tumour
CC cells, including those with a multidrug resistant phenotype. The
CC pharmaceutical composition can be used to treat tumours, specifically
CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
CC cervix, uterus, skin, prostate, liver and colon.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 99; DB 21; Length 13;
Best Local Similarity 100.0%; Pred. No. 9.8e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKWPWWPWRR 13
||||||||||||||
Db 1 ILPWKWPWWPWRR 13

RESULT 10
AAY44666
ID AAY44666 standard; peptide; 13 AA.
XX
AC AAY44666;
XX
DT 18-APR-2000 (first entry)
XX
DE Crosslink-stabilised indolicidin analog Indol 1-13(W6/9).
XX
KW Crosslinked indolicidin analog; X-indolicidin; Indol 1-13(W6/9);
KW stability; bovine neutrophil; antimicrobial; antibacterial; fungicide;
KW protozoacide; virucide; anti-HIV; human immunodeficiency virus-1;
KW HIV-1; gram positive bacteria; gram negative; Staphylococcus aureus;
KW Escherichia coli; Salmonella typhimurium; yeast; fungi; protozoa;
KW Candida albicans; Cryptococcus neoformans; Giardia; Acanthamoeba.
XX
OS Synthetic.
OS Bos sp.
XX
FH Key Location/Qualifiers
FT Modified-site 6..9
FT /note= "Residues at positions 6 and 9 form a
FT di-tryptophan crosslink"
FT Modified-site 13
FT /note= "C-terminal amide"
XX
PN WO9965510-A1.
XX
PD 23-DEC-1999.
XX
PF 20-MAY-1999; 99WO-US11165.
XX
PR 18-JUN-1998; 98US-0099631.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Selsted ME, Osapay K;
XX
DR WPI; 2000-147133/13.
XX
PT Crosslinked indolicidin analogs with antimicrobial activity against
PT bacteria, yeast, fungi, protozoa and viruses -
XX
PS Claim 3; Page 39; 53pp; English.
XX
CC The patent discloses crosslinked analogs of indolicidin (Indol 1-13)
CC which is a naturally occurring peptide isolated from bovine neutrophils
CC and has antimicrobial activity. The crosslinked indolicidin
CC (X-indolicidin) analogs are stable and have antimicrobial activity
CC against gram positive and negative bacteria (e.g. Staphylococcus aureus,

CC Escherichia coli and Salmonella typhimurium), yeasts and fungi (e.g.
CC Candida albicans, Cryptococcus neoformans), protozoa (e.g. Giardia
CC species and Acanthamoeba species), and viruses (e.g. HIV-1).
CC They can be used for reducing or inhibiting the growth or survival of
CC microorganisms in an environment e.g. a food or food product, a
CC solution, an inanimate object comprising a surface, or a mammal.
CC The present sequence is a specifically claimed X-indolicidin
CC analog, Indol 1-13(W6/9) which contains a di-tryptophan crosslink.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 99; DB 21; Length 13;
Best Local Similarity 100.0%; Pred. No. 9.8e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKWPWWPWRR 13
||||||||||||||
Db 1 ILPWKWPWWPWRR 13

RESULT 11
AAY44324
ID AAY44324 standard; peptide; 13 AA.
XX
AC AAY44324;
XX
DT 29-FEB-2000 (first entry)
XX
DE Antimicrobial peptide, Indolicidin.
XX
KW purF gene; glutamine pyrophosphoribosyl pyrophosphate amidotransferase;
KW purF derivative; fusion partner; antimicrobial peptide; Indolicidin;
KW mass production; cleavage site; hydroxylamine; CNBr; DNA construct; cow;
KW neutralise; toxicity; pharmaceutical industry; food industry.
XX
OS Bos taurus.
XX
PN WO9964611-A1.
XX
PD 16-DEC-1999.
XX
PF 08-JUN-1999; 99WO-KR00282.
XX
PR 09-JUN-1998; 98KR-0022117.
PR 14-MAY-1999; 99KR-0017920.
XX
PA (SAMY-) SAMYANG GENEX CORP.
XX
PI Kim JH, Kang MH, Lee J, Park SH, Lee JW, Hong SS, Lee H;
XX
DR WPI; 2000-097542/08.
DR N-PSDB; AAZ44324.
XX
PT New DNA constructs useful for mass production of antimicrobial peptides
PT in microorganism hosts -
XX
PS Claim 1; Fig 1; 67pp; English.
XX
CC The present amino acid sequence is an antimicrobial peptide, Indolicidin
CC derived from cow, Bos taurus. It is used along with a
CC derivative of purF gene sequence that functions as a fusion partner.
CC A DNA construct that comprises, this antimicrobial peptide encoding
CC sequence and the entire, partial or derivative of purF gene, is used for
CC mass production of the antimicrobial peptide in microorganisms without
CC killing the host cells. Use of the purF gene derivative sequence,
CC neutralises the toxicity of the antimicrobial peptides against the host
CC microorganism. The antimicrobial peptides are useful commercially in the
CC pharmaceutical and food industries.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 99; DB 21; Length 13;
Best Local Similarity 100.0%; Pred. No. 9.8e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKWPWWPWRR 13
||||||||||||||
Db 1 ILPWKWPWWPWRR 13

RESULT 12
AAY55056
ID AAY55056 standard; peptide; 13 AA.
XX
AC AAY55056;
XX
DT 23-FEB-2000 (first entry)
XX
DE Non-amidated indolicidin peptide.
XX
KW Indolicidin; bactenecin; sulphate-reducing bacteria; growth inhibitor;
KW corrosion; degradation; metal; concrete; cement; dental implant; biofilm.
XX
OS Bacillus sp.
XX
PN WO9956553-A1.
XX
PD 11-NOV-1999.
XX
PF 03-MAY-1999; 99WO-US09675.
XX
PR 06-MAY-1998; 98US-0074037.
PR 31-MAR-1999; 99US-0282277.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Wood TK, Jayaraman A, Earthman JC;
XX
DR WPI; 2000-052882/04.
XX
PT Inhibiting growth of sulphate-reducing bacteria using other bacteria,
PT particularly for protection of metals and concrete -
XX
PS Example 4; Page 41; 84pp; English.
XX
CC This sequence represents the non-amidated indolicidin peptide.
CC The invention relates to a method for inhibiting growth of
CC sulphate-reducing bacteria (A) on a material (B) sensitive to corrosion
CC or degradation, by applying to (B) a bacterium (C) that secretes a
CC compound (I) able to inhibit growth of (A). The method is used to protect
CC metal, concrete or cement against corrosion and degradation, but (B) can
CC also be used to protect dental implants. (B) is present in an open or
CC closed system (e.g. water cooling tower, liquid storage container, fuel
CC tank, sewer or drainage system etc.) or part of a bridge or other
CC structure. The method is more effective and less expensive than known
CC methods for inhibiting (A), and reduces the amount of toxic chemicals
CC released. Conventional biofilms of aerobic organisms tend to encourage
CC growth of (A), and addition of (C) to the biofilm prevents this. A
CC single application of (C) lasts for a long time, and (I) are produced
CC exactly where they are required and inhibit (A) without significant
CC impact on other organisms (this effect includes reducing resistance of
CC (A) to conventional biocides, which may then be used in reduced
CC amounts). If local damage to the biofilm occurs, the underlying
CC material is still protected by diffusion of (I) from neighbouring areas.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 99; DB 21; Length 13;
Best Local Similarity 100.0%; Pred. No. 9.8e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKWPWWPWRR 13
||||||||||||||
Db 1 ILPWKWPWWPWRR 13

```
RESULT 13
AAY57123
ID   AAY57123 standard; peptide; 13 AA.
XX
AC   AAY57123;
XX
DT   28-FEB-2000  (first entry)
XX
DE   Naturally occurring bovine indolicidin peptide Indol 1-13.
XX
KW   Indolicidin analogue; antimicrobial activity; helminth; bacteria; virus;
KW   treatment; inhibit growth; micro-organism; contact lens solution;
KW   transgenic plant; surgical instrument; yeast; fungi; protozoa.
XX
OS   Bos sp.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   13
FT                   /note= "C-terminal amide"
XX
PN   WO9958141-A1.
XX
PD   18-NOV-1999.
XX
PF   05-MAY-1999;   99WO-US09942.
XX
PR   12-MAY-1998;   98US-0076227.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Selsted ME;
XX
DR   WPI; 2000-053028/04.
XX
PT   New indolicidin analogues, active against bacteria, yeast, fungi,
PT   protozoa and virus, used for, e.g. treating infections -
XX
PS   Example 1; Page 28; 62pp; English.
XX
CC   This sequence is a naturally occurring indolicidin peptide.
CC   Peptides AAY57109-Y57138 and AAY57143-Y57144s are new indolicidin
CC   analogues, which have a homoserine residue and/or a truncated amino
CC   terminal region. The analogues have the following amino acid sequence:
CC   Xaa1-Xaa2-Xaa3-Xaa4-Xaa5-Xaa6-Pro-Xaa6-Xaa6-Pro-Xaa6-Xaa7-Xaa7-Xaa8
CC   Where:
CC   Xaa1 = Ile, Leu, Val, Ala, Gly or absent;
CC   Xaa2 = Ile, Leu, Val, Ala, Gly or absent;
CC   Xaa3 = Pro or absent;
CC   Xaa4 = Trp, Phe or absent;
CC   Xaa5 = Arg, Lys or absent;
CC   Xaa6 = Trp or Phe;
CC   Xaa7 = Arg, Lys or absent;
CC   Xaa8 = homoserine (Hse), Met, Met-Xaa9-Met or absent, and
CC   Xaa9 = at least one amino acid;
CC   provided that if Xaa1 is present, Xaa8 = Hse, Met or Met-Xaa9-Met;
CC   and further provided that: if Xaa2 is absent, Xaa1 is absent; if Xaa3 is
CC   absent, Xaa1 and Xaa2 are absent; if Xaa4 is absent, Xaa1, Xaa2 and Xaa3
CC   are absent; and if Xaa5 is absent, Xaa1, Xaa2, Xaa3 and Xaa4 are absent.
CC   The indolicidin analogues can be used to create a fusion polypeptide
CC   consisting of the analogue linked to a peptide. The indolicidin
CC   analogues have antimicrobial activity against gram positive bacteria,
CC   gram negative bacteria, yeast, fungus, protozoa and viruses (e.g. HIV-1).
CC   They are also active against helminths. The analogues can be used for
CC   reducing or inhibiting growth or survival of a microorganism. They can be
CC   used for treating infections. They can also be included in a liquid such
CC   as water or an aqueous solution, e.g. contact lens solution. The
CC   analogues have potential uses in food products, and in objects such as
CC   the surface of an instrument used to prepare food or to perform surgery.
CC   Transgenic plants or animals useful in the food industry can be produced
CC   by introducing a nucleic acid molecule encoding an indolicidin analogue
CC   into the germline cells of such organisms.
XX
SQ   Sequence   13 AA;

  Query Match            100.0%;  Score 99;  DB 21;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 9.8e-07;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPWRR 13
              |||||||||||||
Db          1 ILPWKWPWWPWRR 13


RESULT 14
ABP60382
ID   ABP60382 standard; peptide; 13 AA.
XX
AC   ABP60382;
XX
DT   28-MAR-2003  (first entry)
XX
DE   Indolicidin peptide SEQ ID NO 1.
XX
KW   Indolicidin; ophthalmic; disinfection; contact lens; antimicrobial;
KW   Pseudomonas aeruginosa; Staphylococcus aureus; Serratia marcescens;
KW   Candida albicans; Fusarium solani.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   13
FT                   /note= "C-terminal CONH2"
XX
PN   WO200071175-A1.
XX
PD   30-NOV-2000.
XX
PF   23-MAY-2000; 2000WO-US14608.
XX
PR   25-MAY-1999;   99US-0318195.
XX
PA   (LARG-) LARGE SCALE BIOLOGY CORP.
PA   (STRI ) SRI INT.
PA   (REGC ) UNIV CALIFORNIA.
PA   (WESL-) WESLEY-JESSEN CORP.
PA   (TUSE/) TUSE D.
PA   (MORT/) MORTELMANS K.
PA   (HOKA/) HOKAMA L A.
PA   (SELS/) SELSTED M E.
PA   (CHAP/) CHAPOY L L.
PA   (QUIN/) QUINN M H.
XX
PI   Tuse D,  Mortelmans K,  Hokama LA,  Selsted ME,  Chapoy LL,  Quinn MH;
XX
DR   WPI; 2001-080322/09.
XX
PT   Ophthalmic composition for storing, cleaning, or disinfecting contact
PT   lens, comprises indolicidin, and buffer having specified halide ion
PT   concentration or Good's buffer  -
XX
PS   Claim 15; Page 68; 91pp; English.
XX
CC   The invention relates to an ophthalmic composition (I) for storing,
CC   cleaning, or disinfecting a contact lens, comprising an indolicidin
CC   antimicrobial peptide and a buffer having a halide ion concentration less
CC   than 0.85 weight%, based on the total weight of (I) or Good's buffer. (I)
CC   is a multipurpose solution for care of a contact lens and is suitable for
CC   contact lens disinfection, storage, cleaning, conditioning, rehydrating,
CC   moistening and lubricating. (I) is useful for disinfecting the contact
CC   lens or contact lens storage vessel such as contact lens vial, contact
CC   lens case or a contact lens shipping package by contacting the lens or
CC   vessel with a disinfecting solution comprising (I). (I) is useful for
CC   packaging a contact lens involving sealing the lens in a container with
CC   (I), where the contact lens is not autoclaved. (I) reduces the number of
CC   Pseudomonas aeruginosa, Staphylococcus aureus and Serratia marcescens
CC   organisms by 3.0 logs or more within 4 hours and the number of Candida
```

CC albicans and Fusarium solani by 1.0 log or more within 18 hours. (I) is
CC self-preserving and requires no additional preservatives or
CC disinfectants. Since indolicidins are safe for topical application to the
CC eye, (I) enables immediate application of the contact lens to the eye
CC without the need for neutralisation, deactivation or washing any of the
CC components of (I). The present sequence is that of an indolicidin peptide
CC of the invention.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 99; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 9.8e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKWPWWPWRR 13
||||||||||||||
Db 1 ILPWKWPWWPWRR 13

RESULT 15
ABP60383
ID ABP60383 standard; peptide; 13 AA.
XX
AC ABP60383;
XX
DT 28-MAR-2003 (first entry)
XX
DE Indolicidin peptide SEQ ID NO 2.
XX
KW Indolicidin; ophthalmic; disinfection; contact lens; antimicrobial;
KW Pseudomonas aeruginosa; Staphylococcus aureus; Serratia marcescens;
KW Candida albicans; Fusarium solani.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Modified-site 13
FT /note= "C-terminal OH"
XX
PN WO200071175-A1.
XX
PD 30-NOV-2000.
XX
PF 23-MAY-2000; 2000WO-US14608.
XX
PR 25-MAY-1999; 99US-0318195.
XX
PA (LARG-) LARGE SCALE BIOLOGY CORP.
PA (STRI) SRI INT.
PA (REGC) UNIV CALIFORNIA.
PA (WESL-) WESLEY-JESSEN CORP.
PA (TUSE/) TUSE D.
PA (MORT/) MORTELMANS K.
PA (HOKA/) HOKAMA L A.
PA (SELS/) SELSTED M E.
PA (CHAP/) CHAPOY L L.
PA (QUIN/) QUINN M H.
XX
PI Tuse D, Mortelmans K, Hokama LA, Selsted ME, Chapoy LL, Quinn MH;
XX
DR WPI; 2001-080322/09.
XX
PT Ophthalmic composition for storing, cleaning, or disinfecting contact
PT lens, comprises indolicidin, and buffer having specified halide ion
PT concentration or Good's buffer
XX
PS Disclosure; Page 16; 91pp; English.
XX
CC The invention relates to an ophthalmic composition (I) for storing,
CC cleaning, or disinfecting a contact lens, comprising an indolicidin
CC antimicrobial peptide and a buffer having a halide ion concentration less
CC than 0.85 weight%, based on the total weight of (I) or Good's buffer. (I)
CC is a multipurpose solution for care of a contact lens and is suitable for
CC contact lens disinfection, storage, cleaning, conditioning, rehydrating,
CC moistening and lubricating. (I) is useful for disinfecting the contact
CC lens or contact lens storage vessel such as contact lens vial, contact
CC lens case or a contact lens shipping package by contacting the lens or
CC vessel with a disinfecting solution comprising (I). (I) is useful for
CC packaging a contact lens involving sealing the lens in a container with
CC (I), where the contact lens is not autoclaved. (I) reduces the number of
CC Pseudomonas aeruginosa, Staphylococcus aureus and Serratia marcescens
CC organisms by 3.0 logs or more within 4 hours and the number of Candida
CC albicans and Fusarium solani by 1.0 log or more within 18 hours. (I) is
CC self-preserving and requires no additional preservatives or
CC disinfectants. Since indolicidins are safe for topical application to the
CC eye, (I) enables immediate application of the contact lens to the eye
CC without the need for neutralisation, deactivation or washing any of the
CC components of (I). The present sequence is that of an indolicidin peptide
CC of the invention.
XX
SQ Sequence 13 AA;

Query Match 100.0%; Score 99; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 9.8e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKWPWWPWRR 13
||||||||||||||
Db 1 ILPWKWPWWPWRR 13

Search completed: October 2, 2003, 10:01:28
Job time : 78 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 2, 2003, 09:59:37 ; Search time 41 Seconds
(without alignments)
30.493 Million cell updates/sec

Title: US-09-444-281-85
Perfect score: 99
Sequence: 1 ILPWKWPWWPWRR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	99	100.0	144	1	JC1222	indolicidin precur
2	54	54.5	1173	1	VGIHHC	E2 glycoprotein pr
3	53.5	54.0	299	2	T12505	hypothetical prote
4	53	53.5	327	2	E72851	AcOrf-13 protein -
5	53	53.5	331	2	T41758	AcMNPV orf13 - Bom
6	51	51.5	55	2	E90626	ATP synthase F0 ch
7	51	51.5	689	2	AC1927	hypothetical prote
8	51	51.5	1038	2	I38935	bone morphogenetic
9	50.5	51.0	970	2	T28234	ORF MSV076 probabl
10	50	50.5	83	2	B72392	hypothetical prote
11	50	50.5	337	2	G95922	probable glycosylt
12	50	50.5	498	1	JT0751	ferredoxin-NADP re
13	49.5	50.0	296	2	T03562	conserved hypothet
14	49	49.5	60	2	A56547	sex-peptide precur
15	49	49.5	425	2	E84631	probable serine ca
16	49	49.5	467	2	E89605	protein F18G5.2 [i
17	48.5	49.0	111	2	T29295	hypothetical prote
18	48	48.5	55	2	T11105	H+-transporting tw
19	48	48.5	265	2	AH0755	conserved hypothet
20	48	48.5	400	2	AF2107	hypothetical prote
21	47	47.5	55	1	PWXL8	H+-transporting tw
22	47	47.5	55	2	S68132	H+-transporting tw
23	47	47.5	55	2	S08424	H+-transporting tw
24	47	47.5	55	2	E90618	ATP synthase F0 ch
25	47	47.5	55	2	T11538	H+-transporting tw
26	47	47.5	55	2	T11184	H+-transporting tw
27	47	47.5	55	2	T11291	H+-transporting tw
28	47	47.5	55	2	T09861	H+-transporting tw
29	47	47.5	55	2	T09951	H+-transporting tw
30	47	47.5	55	2	T11768	H+-transporting tw
31	47	47.5	55	2	T11304	H+-transporting tw
32	47	47.5	248	2	S23449	NADH oxidase (H2O2
33	47	47.5	253	2	G70715	hypothetical prote
34	47	47.5	297	2	D87260	integral membrane
35	47	47.5	456	2	T18963	hypothetical prote
36	47	47.5	496	2	A54770	N-acetylglucosamin
37	47	47.5	534	1	S75101	hypothetical prote
38	47	47.5	728	2	T51071	related to trfA pr
39	47	47.5	1112	2	S70522	cyclic nucleotide
40	46.5	47.0	1299	2	AB2244	hypothetical prote
41	46	46.5	54	1	S04619	H+-transporting tw
42	46	46.5	55	2	T11171	H+-transporting tw
43	46	46.5	55	2	T12413	H+-transporting tw
44	46	46.5	55	2	E58892	H+-transporting tw
45	46	46.5	55	2	E90612	ATP synthase F0 ch

ALIGNMENTS

RESULT 1
JC1222
indolicidin precursor - bovine
N;Alternate names: antimicrobial peptide
C;Species: Bos primigenius taurus (cattle)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: JC1222; A42387; S25664
R;del Sal, G.; Storici, P.; Schneider, C.; Romeo, D.; Zanetti, M.
Biochem. Biophys. Res. Commun. 187, 467-472, 1992
A;Title: cDNA cloning of the neutrophil bactericidal peptide indolicidin.
A;Reference number: JC1222; MUID:92392368; PMID:1520337
A;Accession: JC1222
A;Molecule type: mRNA
A;Residues: 1-144 <SAL>
A;Cross-references: EMBL:X67340; NID:g462; PIDN:CAA47755.1; PID:g463
A;Experimental source: bone marrow
R;Selsted, M.E.; Novotny, M.J.; Morris, W.L.; Tang, Y.Q.; Smith, W.; Cullor, J.S.
J. Biol. Chem. 267, 4292-4295, 1992
A;Title: Indolicidin, a novel bactericidal tridecapeptide amide from neutrophils.
A;Reference number: A42387; MUID:92165771; PMID:1537821
A;Accession: A42387
A;Molecule type: protein
A;Residues: 131-143 <SEL>
A;Experimental source: neutrophils
A;Note: sequence extracted from NCBI backbone (NCBIP:83840)
C;Superfamily: cathelin; cystatin homology
C;Keywords: amidated carboxyl end
F;1-29/Domain: signal sequence #status predicted <SIG>
F;22-129/Domain: cystatin homology <CYS>
F;30-130/Domain: propeptide #status predicted <PRO>
F;131-143/Product: indolicidin #status experimental <MAT>
F;143/Modified site: amidated carboxyl end (Arg) (amide in mature form from following

Query Match 100.0%; Score 99; DB 1; Length 144;
Best Local Similarity 100.0%; Pred. No. 7.4e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ILPWKWPWWPWRR 13
              |||||||||||||
Db        131 ILPWKWPWWPWRR 143
```

RESULT 2
VGIHHC
E2 glycoprotein precursor - human coronavirus (strain 229E)
N;Alternate names: peplomer glycoprotein; S glycoprotein; spike glycoprotein
C;Species: human coronavirus
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jun-2000
C;Accession: A34766; S05460
R;Raabe, T.; Schelle-Prinz, B.; Siddell, S.G.
J. Gen. Virol. 71, 1065-1073, 1990

```
A;Title: Nucleotide sequence of the gene encoding the spike glycoprotein of human corona
A;Reference number: A34766; MUID:90264837; PMID:2345367
A;Accession: A34766
A;Molecule type: mRNA
A;Residues: 1-1173 <RAA>
A;Cross-references: EMBL:X16816; NID:g58926; PIDN:CAA34723.1; PID:g58927
A;Experimental source: strain 229E
R;Raabe, T.; Siddell, S.
Nucleic Acids Res. 17, 6387, 1989
A;Title: Nucleotide sequence of the human coronavirus HCV 229E mRNA 4 and mRNA 5 unique
A;Reference number: A34038; MUID:89366667; PMID:2701946
A;Accession: S05460
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1159-1173 <RA2>
A;Cross-references: EMBL:X15654; NID:g58921; PIDN:CAA33680.1; PID:g1334827
C;Superfamily: coronavirus E2 glycoprotein
C;Keywords: glycoprotein; transmembrane protein
F;1-15/Domain: signal sequence #status predicted <SIG>
F;16-1173/Product: E2 glycoprotein #status predicted <MAT>
F;1116-1138/Domain: transmembrane #status predicted <TMN>
F;23,62,98,147,171,176,220,243,326,333,440,464,518,538,542,568,581,587,663,671,930,1015,

  Query Match             54.5%;  Score 54;  DB 1;  Length 1173;
  Best Local Similarity   85.7%;  Pred. No. 21;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 KWPWWPW 11
              ||||| |
Db       1113 KWPWWVW 1119


RESULT 3
T12505
hypothetical protein DKFZp434C192.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 23-Jul-1999
C;Accession: T12505
R;Ansorge, W.; Wirkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z17527
A;Accession: T12505
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-299 <ANS>
A;Cross-references: EMBL:AL096753
A;Experimental source: adult testis; clone DKFZp434C192
C;Genetics:
A;Note: DKFZp434C192.1

  Query Match             54.0%;  Score 53.5;  DB 2;  Length 299;
  Best Local Similarity   57.1%;  Pred. No. 6.4;
  Matches    8;  Conservative    0;  Mismatches    3;  Indels    3;  Gaps    1;

Qy          3 PW---KWPWWPWRR 13
              ||     ||||| |
Db         30 PWYGSASPWWPWSR 43


RESULT 4
E72851
AcOrf-13 protein - Autographa californica nuclear polyhedrosis virus
C;Species: Autographa californica nuclear polyhedrosis virus, AcMNPV
A;Note: dsDNA virus
C;Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 12-Nov-1999
C;Accession: E72851
R;Ayres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R.D.
Virology 202, 586-605, 1994
A;Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis virus.
A;Reference number: A72850; MUID:94303173; PMID:8030224
A;Accession: E72851
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-327 <AYR>
A;Cross-references: GB:L22858; NID:g510708; PIDN:AAA66643.1; PID:g559082
C;Genetics:
A;Gene: AcOrf-13

  Query Match             53.5%;  Score 53;  DB 2;  Length 327;
  Best Local Similarity   54.5%;  Pred. No. 8;
  Matches    6;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPW 11
              :| | | || |
Db          1 MLSWLWNWWMW 11


RESULT 5
T41758
AcMNPV orf13 - Bombyx mori nuclear polyhedrosis virus (isolate T3)
C;Species: Bombyx mori nuclear polyhedrosis virus, BmSNPV
A;Variety: isolate T3
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C;Accession: T41758
R;Gomi, S.; Majima, K.; Maeda, S.
J. Gen. Virol. 80, 1323-1337, 1999
A;Title: Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus.
A;Reference number: Z22020; MUID:99281911; PMID:10355780
A;Accession: T41758
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-331 <KAM>
A;Cross-references: EMBL:L33180; NID:g3745835; PIDN:AAC63687.1; PID:g3745840
A;Experimental source: isolate T3
C;Genetics:
A;Note: Orf_5

  Query Match             53.5%;  Score 53;  DB 2;  Length 331;
  Best Local Similarity   54.5%;  Pred. No. 8.1;
  Matches    6;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPW 11
              :| | | || |
Db          1 MLSWLWNWWMW 11


RESULT 6
E90626
ATP synthase F0 chain 8 [imported] - Eudromia elegans mitochondrion
C;Species: mitochondrion Eudromia elegans
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 17-May-2002
C;Accession: E90626
R;Haddrath, O.; Baker, A.J.
Proc. R. Soc. Lond. B Biol. Sci. 268, 939-945, 2001
A;Title: Complete mitochondrial DNA geonome sequences of extinct birds: ratite phylog
A;Reference number: A99613; MUID:21263106; PMID:11370967
A;Accession: E90626
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-55 <KUR>
A;Cross-references: GB:NC_002772; NID:g14141818; PIDN:NP_115277.1; GSPDB:GN00163
C;Genetics:
A;Gene: ATP8
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: H+-transporting ATP synthase protein 8
C;Keywords: mitochondrion

  Query Match             51.5%;  Score 51;  DB 2;  Length 55;
  Best Local Similarity   85.7%;  Pred. No. 2.5;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 LPWKWPW 8
              ||| |||
Db         48 LPWSWPW 54
```

```
RESULT 7
AC1927
hypothetical protein all0966 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AC1927
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AC1927
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-689 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB72923.1; PID:g17130312; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all0966

  Query Match             51.5%;  Score 51;  DB 2;  Length 689;
  Best Local Similarity   55.6%;  Pred. No. 29;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 WKWPWWPWR 12
              | | :|||:
Db        141 WHWGYWPWQ 149


RESULT 8
I38935
bone morphogenetic protein receptor II precursor - human
N;Alternate names: activin receptor-like kinase type II; bone morphogenetic protein 4 re
N;Contains: protein kinase (EC 2.7.1.37)
C;Species: Homo sapiens (man)
C;Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 24-Sep-1999
C;Accession: I38935; I55438; I37209
R;Kawabata, M.; Chytil, A.; Moses, H.L.
J. Biol. Chem. 270, 5625-5630, 1995
A;Title: Cloning of a novel type II serine/threonine kinase receptor through interaction
A;Reference number: A55947; MUID:95197572; PMID:7890683
A;Accession: I38935
A;Molecule type: mRNA
A;Residues: 1-1038 <KAW>
A;Cross-references: EMBL:U20165; NID:g704361; PIDN:AAC50105.1; PID:g704362
R;Nohno, T.; Ishikawa, T.; Saito, T.; Hosokawa, K.; Noji, S.; Wolsing, D.H.; Rosenbaum,
J. Biol. Chem. 270, 22522-22526, 1995
A;Title: Identification of a human type II receptor for bone morphogenetic protein-4 tha
A;Reference number: I55438; MUID:95403457; PMID:7673243
A;Accession: I55438
A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMB
A;Molecule type: mRNA
A;Residues: 1-1038 <NOH>
A;Cross-references: GB:D50516; NID:g807712; PIDN:BAA09094.1; PID:g807713
R;Rosenzweig, B.L.; Imamura, T.; Okadome, T.; Cox, G.N.; Yamashita, H.; ten Dijke, P.; H
Proc. Natl. Acad. Sci. U.S.A. 92, 7632-7636, 1995
A;Title: Cloning and characterization of a human type II receptor for bone morphogenetic
A;Reference number: I37209; MUID:95372334; PMID:7644468
A;Accession: I37209
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-827,'R',829-1038 <ROS>
A;Cross-references: EMBL:Z48923; NID:g1009409; PIDN:CAA88759.1; PID:g1009410
C;Genetics:
A;Gene: GDB:BMPR2; BRK-3; T-ALK; BMPR3; BMPR-II
A;Cross-references: GDB:642243; OMIM:600799
A;Map position: 20pter-20qter
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C;Keywords: ATP; glycoprotein; phosphotransferase; receptor; transmembrane protein
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-1038/Product: bone morphogenetic protein receptor II #status predicted <MAT>
F;27-150/Domain: extracellular #status predicted <EXT>
F;151-170/Domain: transmembrane #status predicted <TRM>
F;201-508/Domain: protein kinase homology <KIN>
F;209-217/Region: protein kinase ATP-binding motif
F;55,110,126/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             51.5%;  Score 51;  DB 2;  Length 1038;
  Best Local Similarity   66.7%;  Pred. No. 44;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 PWKWPWWPW 11
              ||: || ||
Db          8 PWRVPWLPW 16


RESULT 9
T28234
ORF MSV076 probable spheroidin - Melanoplus sanguinipes entomopoxvirus
C;Species: Melanoplus sanguinipes entomopoxvirus
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C;Accession: T28234
R;Afonso, C.L.; Tulman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L.
J. Virol. 73, 533-552, 1999
A;Title: The genome of Melanoplus sanguinipes entomopoxvirus.
A;Reference number: Z20484; MUID:99102612; PMID:9847359
A;Accession: T28234
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-970 <AFO>
A;Cross-references: EMBL:AF063866; NID:g4049647; PIDN:AAC97813.1; PID:g4049853
C;Genetics:
A;Note: MSV076

  Query Match             51.0%;  Score 50.5;  DB 2;  Length 970;
  Best Local Similarity   57.1%;  Pred. No. 47;
  Matches    8;  Conservative    3;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          1 ILPWKWPW-WPWRR 13
              |||: :|| ||: |
Db        838 ILPYPYPWPWPYNR 851


RESULT 10
B72392
hypothetical protein - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: B72392
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hic
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
 C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: B72392
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-83 <ARN>
A;Cross-references: GB:AE001713; GB:AE000512; NID:g4980809; PIDN:AAD35403.1; PID:g498
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM0315

  Query Match             50.5%;  Score 50;  DB 2;  Length 83;
  Best Local Similarity   62.5%;  Pred. No. 4.9;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 WKWPWWPW 11
              | | :|||
Db          7 WSWGFWPW 14


RESULT 11
```

```
G95922
probable glycosyltransferase protein SMb21068 [imported] - Sinorhizobium meliloti (strai
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C;Accession: G95922
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: G95922
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-337 <KUR>
A;Cross-references: GB:AL591985; PIDN:CAC49047.1; PID:g15140532; GSPDB:GN00167
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMb21068
A;Genome: plasmid

  Query Match             50.5%;  Score 50;  DB 2;  Length 337;
  Best Local Similarity   66.7%;  Pred. No. 19;
  Matches    6;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 PWKWPWWPW 11
              ||  | |||
Db        262 PWNTPGWPW 270


RESULT 12
JT0751
ferredoxin-NADP reductase (EC 1.18.1.2), long form precursor - bovine
N;Alternate names: adrenodoxin reductase
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Jul-1994 #sequence_revision 18-Oct-1996 #text_change 03-Jun-2002
C;Accession: JT0751; JT0079; JS0390; S03558; PS0003; A29604; S52100
R;Takata, Y.; Sagara, Y.; Kono, A.; Sekimizu, K.; Horiuchi, T.
Biol. Pharm. Bull. 16, 1200-1206, 1993
A;Title: Gene structure of bovine adrenodoxin reductase.
A;Reference number: JT0751; MUID:94177140; PMID:8130767
A;Accession: JT0751
A;Molecule type: DNA
A;Residues: 1-498 <TAK>
A;Cross-references: GB:D83475; NID:g1199916; PIDN:BAA11921.1; PID:g4521308
A;Experimental source: adrenal cortex
A;Note: the authors translated the codon GTC for residue 205 as Gly
R;Sagara, Y.; Takata, Y.; Miyata, T.; Hara, T.; Horiuchi, T.
J. Biochem. 102, 1333-1336, 1987
A;Title: Cloning and sequence analysis of adrenodoxin reductase cDNA from bovine adrenal
A;Reference number: JT0079; MUID:88198050; PMID:3448086
A;Accession: JT0079
A;Molecule type: mRNA
A;Residues: 1-204,211-498 <SAG>
A;Cross-references: GB:D00211; NID:g217433; PIDN:BAA00150.1; PID:g217434
A;Note: the deduced sequence is partially confirmed by amino acid sequencing of 15 isola
R;Sagara, Y.
submitted to DDBJ, September 1989
A;Reference number: JS0390
A;Contents: revision, insertion of residues 205-210
A;Accession: JS0390
A;Molecule type: mRNA
A;Residues: 56-498 <SA2>
R;Hanukoglu, I.; Gutfinger, T.
Eur. J. Biochem. 180, 479-484, 1989
A;Title: cDNA sequence of adrenodoxin reductase. Identification of NADP-binding sites in
A;Reference number: S03558; MUID:89170752; PMID:2924777
A;Accession: S03558
A;Molecule type: mRNA
A;Residues: 155-204,211-498 <HAN>
A;Cross-references: EMBL:X13736; NID:g65; PIDN:CAA32002.1; PID:g833776
A;Note: 405-Ser was also found
R;Hamamoto, I.; Kurokohchi, K.; Tanaka, S.; Ichikawa, Y.
Biochim. Biophys. Acta 953, 207-213, 1988
A;Title: Adrenoferredoxin-binding peptide of NADPH-adrenoferredoxin reductase.
A;Reference number: PS0003; MUID:88184054; PMID:3355838
A;Accession: PS0003
A;Molecule type: protein
A;Residues: 33-41,'S',43-62;260-283,'TM';496-498 <HAM>
A;Note: a cyanogen bromide peptide binds to adrenoferredoxin
R;Nonaka, Y.; Murakami, H.; Yabusaki, Y.; Kuramitsu, S.; Kagamiyama, H.; Yamano, T.;
Biochem. Biophys. Res. Commun. 145, 1239-1247, 1987
A;Title: Molecular cloning and sequence analysis of full-length cDNA for mRNA of adre
A;Reference number: A29604; MUID:87270696; PMID:3038094
A;Accession: A29604
A;Molecule type: mRNA
A;Residues: 1-76,'R',78-80,'VWLALTTPRSRMLL',95-123,'RVYRLT',129-204,211-273,'R',275-3
A;Cross-references: GB:M17029; NID:g162628; PIDN:AAA30362.1; PID:g162629
A;Experimental source: adrenal cortex
R;Warburton, R.J.; Seybert, D.W.
Biochim. Biophys. Acta 1246, 39-46, 1995
A;Title: Structural and functional characterization of bovine adrenodoxin reductase b
A;Reference number: S52100; MUID:95110846; PMID:7811729
A;Accession: S52100
A;Status: preliminary
A;Molecule type: protein
A;Residues: 'X',34-41,'X',43-48,'X',50-51;304-306,'X',308-309,'X',311-326 <WAR>
C;Comment: Ferredoxin-NADP+ reductase is localized in the matrix of adrenal cortex mi
erredoxin-NADP+ reductase, adrenodoxin and two forms of cytochrome P-450.
C;Genetics:
A;Introns: 27/1; 59/3; 91/3; 132/3; 170/3; 204/3; 246/3; 275/1; 341/3; 399/1; 456/1
C;Function:
A;Description: catalyzes the reversible reduction of NADP+ by reduced ferredoxin or r
C;Superfamily: human ferredoxin-NADP+ reductase
C;Keywords: alternative splicing; flavoprotein; mitochondrion; monomer; NADP; oxidore
F;1-32/Domain: transit peptide (mitochondrion) #status predicted <SIG>
F;33-498/Product: ferredoxin-NADP+ reductase, long form #status predicted <MAT>
F;33-204,211-498/Product: ferredoxin-NADP+ reductase, short form #status experimental
F;40-70/Region: beta-alpha-beta FAD nucleotide-binding fold
F;180-190/Region: NADP binding #status predicted
F;281/Binding site: substrate (Lys) #status experimental

  Query Match             50.5%;  Score 50;  DB 1;  Length 498;
  Best Local Similarity   66.7%;  Pred. No. 28;
  Matches    6;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 PWKWPWWPW 11
              |  | ||||
Db          3 PRCWRWWPW 11


RESULT 13
T03562
conserved hypothetical protein - Rhodobacter capsulatus
C;Species: Rhodobacter capsulatus
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 02-Aug-2002
C;Accession: T03562
R;Vlcek, C.; Paces, V.; Maltsev, N.; Paces, J.; Haselkorn, R.; Fonstein, M.
Proc. Natl. Acad. Sci. U.S.A. 94, 9384-9388, 1997
A;Title: Sequence of a 189-kb segment of the chromosome of Rhodobacter capsulatus SB1
A;Reference number: Z14955; MUID:97404404; PMID:9256491
A;Accession: T03562
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-296 <VLC>
A;Cross-references: EMBL:AF010496; NID:g3128256; PIDN:AAC16215.1; PID:g3128363
C;Genetics:
A;Map position: 1
C;Superfamily: hypothetical protein ydeD
```

Query Match 50.0%; Score 49.5; DB 2; Length 296;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 9; Conservative 3; Mismatches 1; Indels 5; Gaps 2;

```
Qy          1 ILPWKWPW----WP-WRR 13
              :||: |||    || :||
Db         52 LLPFAWPWLRRDWPLYRR 69
```

RESULT 14
A56547
sex-peptide precursor - Drosophila suzukii
N;Alternate names: male accessory gland peptide
C;Species: Drosophila suzukii
C;Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 21-Jul-2000
C;Accession: A56547; B56547
R;Schmidt, T.; Choffat, Y.; Schneider, M.; Hunziker, P.; Fuyama, Y.; Kubli, E.
Insect Biochem. Mol. Biol. 23, 571-579, 1993
A;Title: Drosophila suzukii contains a peptide homologous to the Drosophila melanogaster
A;Reference number: A56547; MUID:93357785; PMID:8353518
A;Accession: A56547
A;Molecule type: DNA
A;Residues: 1-60 <SCH>
A;Cross-references: GB:S64573; NID:g409350; PIDN:AAB27872.2; PID:g7548732
A;Note: sequence modified after extraction from NCBI backbone
A;Note: authors translated the codon TGC for residue 12 as Trp
A;Note: sequence extracted from NCBI backbone (NCBIN:136396)
A;Accession: B56547
A;Molecule type: protein
A;Residues: 20-60 <SC2>
A;Note: sequence modified after extraction from NCBI backbone
C;Genetics:
A;Gene: FlyBase:Dsuz/SP
A;Cross-references: FlyBase:FBgn0012991
A;Introns: 44/1
C;Keywords: neuropeptide
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-60/Product: sex-peptide #status experimental <MAT>

Query Match 49.5%; Score 49; DB 2; Length 60;
Best Local Similarity 50.0%; Pred. No. 4.8;
Matches 7; Conservative 1; Mismatches 2; Indels 4; Gaps 1;

```
Qy          4 WKWPW----WPWRR 13
              |:|||     || |
Db         20 WEWPWNKQKKPWER 33
```

RESULT 15
E84631
probable serine carboxypeptidase II [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C;Accession: E84631
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: E84631
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-425 <STO>
A;Cross-references: GB:AE002093; NID:g3738328; PIDN:AAC63669.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g24010
A;Map position: 2
C;Superfamily: serine carboxypeptidase

Query Match 49.5%; Score 49; DB 2; Length 425;
Best Local Similarity 70.0%; Pred. No. 32;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          2 LPWKWPWWPW 11
              || | ||:||
Db        363 LPVKTPWYPW 372
```

Search completed: October 2, 2003, 10:04:46
Job time : 42 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 2, 2003, 09:57:07 ; Search time 25 Seconds
(without alignments)
24.454 Million cell updates/sec

Title: US-09-444-281-85
Perfect score: 99
Sequence: 1 ILPWKWPWWPWRR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	99	100.0	144	1	INDC_BOVIN	P33046 bos taurus
2	54	54.5	1173	1	VGL2_CVH22	P15423 human coron
3	53	53.5	327	1	Y013_NPVAC	P41423 autographa
4	51	51.5	1038	1	BMR2_HUMAN	Q13873 homo sapien
5	50	50.5	55	1	ATP8_CORCN	Q9tbi6 corythaixoi
6	50	50.5	492	1	ADRO_BOVIN	P08165 bos taurus
7	48	48.5	55	1	ATP8_PELSU	O79674 pelomedusa
8	47.5	48.0	279	1	ELO1_HUMAN	Q9bw60 homo sapien
9	47.5	48.0	279	1	ELO1_MOUSE	Q9jlj5 mus musculu
10	47	47.5	55	1	ATP8_GADMO	P15996 gadus morhu
11	47	47.5	55	1	ATP8_ONCMY	P48179 oncorhynchu
12	47	47.5	55	1	ATP8_PRODO	Q35416 protopterus
13	47	47.5	55	1	ATP8_SALAL	Q9xn27 salvelinus
14	47	47.5	55	1	ATP8_SALFO	Q9xn35 salvelinus
15	47	47.5	55	1	ATP8_SCYCA	O79405 scyliorhinu
16	47	47.5	55	1	ATP8_SQUAC	Q9zz50 squalus aca
17	47	47.5	55	1	ATP8_VIRAL	Q9me31 vireo altil
18	47	47.5	55	1	ATP8_XENLA	P03931 xenopus lae
19	47	47.5	253	1	Y945_MYCTU	P71564 mycobacteri
20	47	47.5	1112	1	CN3B_HUMAN	Q13370 homo sapien
21	46	46.5	54	1	ATP8_CHICK	P14093 gallus gall
22	46	46.5	55	1	ATP8_ANAPL	P50655 anas platyr
23	46	46.5	55	1	ATP8_AYTAM	Q9xkz5 aythya amer
24	46	46.5	55	1	ATP8_CHAPE	Q9tbj9 chaetura pe
25	46	46.5	55	1	ATP8_COLPA	Q9mdb7 columbina p
26	46	46.5	55	1	ATP8_CORCR	Q9tbi7 corythaeola
27	46	46.5	55	1	ATP8_COTJA	P50682 coturnix co
28	46	46.5	55	1	ATP8_LATCH	O03168 latimeria c
29	46	46.5	55	1	ATP8_LOXNO	Q9mdj1 loxigilla n
30	46	46.5	55	1	ATP8_MUSVO	Q9tbi8 musophaga v
31	46	46.5	55	1	ATP8_OPIHO	Q9tbi5 opisthocomu
32	46	46.5	55	1	ATP8_RHEAM	O79396 rhea americ
33	46	46.5	55	1	ATP8_STRCA	O21401 struthio ca
34	45.5	46.0	224	1	Y295_MYCPU	Q98qr6 mycoplasma
35	45	45.5	125	1	YAD3_RHILO	Q98dm8 rhizobium l
36	45	45.5	337	1	BRB1_RAT	P97583 rattus norv
37	45	45.5	433	1	NHG1_PSEPU	P23262 pseudomonas
38	45	45.5	581	1	ACES_BUNFA	Q92035 bungarus fa
39	45	45.5	865	1	E78A_DROME	P45447 drosophila
40	45	45.5	1108	1	CN3B_RAT	Q63085 rattus norv
41	45	45.5	1154	1	VGL2_IBVD2	P12722 avian infec
42	45	45.5	1162	1	VGL2_IBVB	P11223 avian infec
43	45	45.5	1162	1	VGL2_IBVK	P12650 avian infec
44	45	45.5	1162	1	VGL2_IBVM	P12651 avian infec
45	45	45.5	1163	1	VGL2_IBV6	P05135 avian infec

ALIGNMENTS

RESULT 1
INDC_BOVIN
ID INDC_BOVIN STANDARD; PRT; 144 AA.
AC P33046;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Indolicidin precursor.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=92392368; PubMed=1520337;
RA del Sal G., Storici P., Schneider C., Romeo D., Zanetti M.;
RT "cDNA cloning of the neutrophil bactericidal peptide indolicidin.";
RL Biochem. Biophys. Res. Commun. 187:467-472(1992).
RN [2]
RP SEQUENCE OF 131-143.
RC TISSUE=Neutrophils;
RX MEDLINE=92165771; PubMed=1537821;
RA Selsted M.E., Novotny M.J., Morris W.L., Tang Y.-Q., Smith W.,
RA Cullor J.S.;
RT "Indolicidin, a novel bactericidal tridecapeptide amide from
RT neutrophils.";
RL J. Biol. Chem. 267:4292-4295(1992).
CC -!- FUNCTION: POTENT MICROBICIDAL ACTIVITY, ACTIVE AGAINST
CC STAPHYLOCOCCUS AUREUS AND ESCHERICHIA COLI.
CC -!- TISSUE SPECIFICITY: LARGE GRANULES OF NEUTROPHILS.
CC -!- PTM: ELASTASE MIGHT BE RESPONSIBLE FOR ITS MATURATION.
CC -!- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.
CC ----------

CC ----------
DR EMBL; X67340; CAA47755.1; -.
DR PIR; JC1222; JC1222.
DR PDB; 1G89; 17-JAN-01.
DR PDB; 1G8C; 17-JAN-01.
DR PDB; 1HR1; 31-DEC-02.
DR InterPro; IPR001894; Cathelicidin.
DR Pfam; PF00666; Cathelicidins; 1.
DR ProDom; PD001838; Cathelicidin; 1.
DR PROSITE; PS00946; CATHELICIDINS_1; 1.
DR PROSITE; PS00947; CATHELICIDINS_2; 1.
KW Antibiotic; Amidation; Signal; Pyrrolidone carboxylic acid;
KW 3D-structure.
FT SIGNAL 1 29 POTENTIAL.

```
FT   PROPEP       30    130
FT   PEPTIDE     131    143       INDOLICIDIN.
FT   MOD_RES      30     30       PYRROLIDONE CARBOXYLIC ACID (BY
FT                                SIMILARITY).
FT   DISULFID     85     96       BY SIMILARITY.
FT   DISULFID    107    124       BY SIMILARITY.
FT   MOD_RES     143    143       AMIDATION (G-144 PROVIDE AMIDE GROUP).
SQ   SEQUENCE   144 AA;  16479 MW;  E3B1CBBE55C09911 CRC64;

  Query Match             100.0%;  Score 99;  DB 1;  Length 144;
  Best Local Similarity   100.0%;  Pred. No. 5.2e-06;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPWRR 13
              |||||||||||||
Db        131 ILPWKWPWWPWRR 143


RESULT 2
VGL2_CVH22
ID   VGL2_CVH22     STANDARD;      PRT;  1173 AA.
AC   P15423; P89342; P89343; P89344; Q66174; Q990M1; Q990M2; Q990M3;
AC   Q990M4;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   E2 glycoprotein precursor (Spike glycoprotein) (Peplomer protein).
GN   S.
OS   Human coronavirus (strain 229E) (HCoV-229E).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC   Coronaviridae; Coronavirus.
OX   NCBI_TaxID=11137;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90264837; PubMed=2345367;
RA   Raabe T., Schelle-Prinz B., Siddell S.G.;
RT   "Nucleotide sequence of the gene encoding the spike glycoprotein of
RT   human coronavirus HCV 229E.";
RL   J. Gen. Virol. 71:1065-1073(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21262210; PubMed=11369870;
RA   Thiel V., Herold J., Schelle B., Siddell S.G.;
RT   "Infectious RNA transcribed in vitro from a cDNA copy of the human
RT   coronavirus genome cloned in vaccinia virus.";
RL   J. Gen. Virol. 82:1273-1281(2001).
RN   [3]
RP   SEQUENCE OF 98-1113 FROM N.A., AND VARIANTS.
RC   STRAIN=Isolate RW Stock, Isolate P100E, Isolate P11A, and
RC   Isolate P11B;
RA   Bonavia A., Holmes K.V.;
RT   "Viral and cellular changes in a human cell line persistently infected
RT   with human coronavirus HCoV-229E.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 98-1113 FROM N.A., AND VARIANTS.
RC   STRAIN=Isolate ATCC VR-74, Isolate A162, and Isolate LRI 281;
RX   MEDLINE=99086140; PubMed=9870593;
RA   Hays J.P., Myint S.H.;
RT   "PCR sequencing of the spike genes of geographically and
RT   chronologically distinct human coronaviruses 229E.";
RL   J. Virol. Methods 75:179-193(1998).
RN   [5]
RP   SEQUENCE OF 1159-1173 FROM N.A.
RX   MEDLINE=89366667; PubMed=2701946;
RA   Raabe T., Siddell S.;
RT   "Nucleotide sequence of the human coronavirus HCV 229E mRNA 4 and mRNA
RT   5 unique regions.";
RL   Nucleic Acids Res. 17:6387-6387(1989).
RN   [6]
RP   INTERACTION WITH ANPEP.
RX   MEDLINE=22440020; PubMed=12551991;
RA   Bonavia A., Zelus B.D., Wentworth D.E., Talbot P.J., Holmes K.V.;
RT   "Identification of a receptor-binding domain of the spike glycoprotein
RT   of human coronavirus HCoV-229E.";
RL   J. Virol. 77:2530-2538(2003).
RN   [7]
RP   INTERACTION WITH ANPEP.
RX   MEDLINE=22521439; PubMed=12634402;
RA   Breslin J.J., Mork I., Smith M.K., Vogel L.K., Hemmila E.M.,
RA   Bonavia A., Talbot P.J., Sjoestrom H., Noren O., Holmes K.V.;
RT   "Human coronavirus 229E: receptor binding domain and neutralization by
RT   soluble receptor at 37 degrees C.";
RL   J. Virol. 77:4435-4438(2003).
RN   [8]
RP   REVIEW.
RX   MEDLINE=21109095; PubMed=11162792;
RA   Gallagher T.M., Buchmeier M.J.;
RT   "Coronavirus spike proteins in viral entry and pathogenesis.";
RL   Virology 279:371-374(2001).
CC   -!- FUNCTION: Structural protein that makes spikes at the surface of
CC       the virus. Determines enteropathogenicity and virulence of the
CC       virus. Initiates infection by specifically recognizing and binding
CC       the human aminopeptidase ANPEP receptor. Its association with
CC       ANPEP may lead to its conformational change that triggers fusion
CC       between viral and host cellular membrane.
CC   -!- SUBUNIT: Homotrimer. During virus morphogenesis, it is found in a
CC       complex with M and HE proteins (By similarity). Interacts with
CC       ANPEP.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- DOMAIN: The spike S1 domain displays the specificity for the host
CC       receptor.
CC   -!- DOMAIN: The Leucine zipper-like heptad repeats may mediate the
CC       fusion of viral and cellular membranes.
CC   -!- POLYMORPHISM: The strong variation between the different
CC       strains may affect the virulence of the virus.
CC   -!- MISCELLANEOUS: In contrast to serogroup 2, E2 glycoprotein protein
CC       from serogroup 1 is not cleaved.
CC   -!- SIMILARITY: Contains 1 spike S1 domain.
CC   -!- SIMILARITY: Contains 1 spike S2 domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X16816; CAA34723.1; -.
DR   EMBL; AF304460; AAG48592.1; -.
DR   EMBL; AF344186; AAK32188.1; -.
DR   EMBL; AF344187; AAK32189.1; -.
DR   EMBL; AF344188; AAK32190.1; -.
DR   EMBL; AF344189; AAK32191.1; -.
DR   EMBL; Y09923; CAA71056.1; -.
DR   EMBL; Y10051; CAA71146.1; -.
DR   EMBL; Y10052; CAA71147.1; -.
DR   EMBL; X15654; CAA33680.1; -.
DR   PIR; A34766; VGIHHC.
DR   InterPro; IPR002551; Corona_S1.
DR   InterPro; IPR002552; Corona_S2.
DR   Pfam; PF01600; Corona_S1; 1.
DR   Pfam; PF01601; Corona_S2; 1.
KW   Virulence; Glycoprotein; Envelope protein; Transmembrane; Signal;
KW   Coiled coil.
FT   SIGNAL        1     15
FT   CHAIN        16   1173       E2 GLYCOPROTEIN.
FT   DOMAIN       16   1115       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   1116   1135       POTENTIAL.
FT   DOMAIN     1136   1173       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       32    536       SPIKE S1.
FT   DOMAIN      417    547       INTERACTION WITH ANPEP.
FT   DOMAIN      537   1171       SPIKE S2.
FT   DOMAIN     1054   1103       COILED COIL (POTENTIAL).
FT   DOMAIN     1067   1102       LEUCINE ZIPPER-LIKE HEPTAD REPEATS.
```

```
FT   DOMAIN     1136   1157       CYS-RICH.
FT   CARBOHYD     23     23       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     62     62       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     98     98       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    147    147       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    171    171       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    176    176       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    220    220       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    243    243       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    326    326       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    333    333       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    440    440       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    464    464       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    518    518       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    538    538       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    542    542       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    568    568       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    581    581       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    587    587       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    663    663       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    671    671       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    930    930       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1015   1015       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1020   1020       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1037   1037       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1049   1049       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1061   1061       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1066   1066       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1076   1076       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1082   1082       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1096   1096       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT      98     98       N -> S (in isolate LRI 281).
FT   VARIANT     120    120       N -> I (in isolate LRI 281).
FT   VARIANT     127    128       LR -> IS (in isolate A162).
FT   VARIANT     176    176       N -> T (in isolate P100E).
FT   VARIANT     210    210       T -> S (in isolate A162).
FT   VARIANT     223    223       T -> N (in isolate A162).
FT   VARIANT     228    229       DF -> V (in isolate A162).
FT   VARIANT     230    230       C -> L (in isolate LRI 281).
FT   VARIANT     230    230       C -> F (in isolates RW Stock, P11A, P11B,
FT                                 P100E and ATCC VR-74).
FT   VARIANT     248    248       S -> A (in isolate A162).
FT   VARIANT     270    270       D -> Y (in isolate P100E).
FT   VARIANT     295    295       V -> A (in isolate LRI 281).
FT   VARIANT     300    300       T -> M (in isolate P100E).
FT   VARIANT     307    307       D -> N (in isolate A162).
FT   VARIANT     310    311       PQ -> LR (in isolate A162).
FT   VARIANT     314    324       GGKCFNCYPAG -> VGRCYNCRPAV (in isolate
FT                                 A162).
FT   VARIANT     336    336       K -> N (in isolate LRI 281).
FT   VARIANT     349    358       KYVAVYANVG -> QFVGAKFD (in isolate A162).
FT   VARIANT     401    401       V -> M (in isolate A162).
FT   VARIANT     404    411       WAYSKYYT -> LANLNSHN (in isolate A162).
FT   VARIANT     414    414       S -> T (in isolate P100E).
FT   VARIANT     424    424       G -> V (in isolate A162).
FT   VARIANT     430    430       Q -> K (in isolate A162).
FT   VARIANT     441    441       V -> A (in isolate LRI 281).
FT   VARIANT     444    444       D -> N (in isolate A162).
FT   VARIANT     462    462       V -> I (in isolate A162).
FT   VARIANT     481    481       L -> V (in isolate A162).
FT   VARIANT     488    488       K -> N (in isolate A162).
FT   VARIANT     530    530       L -> M (in isolate A162).
FT   VARIANT     577    577       I -> T (in isolate P11A).
FT   VARIANT     578    578       V -> G (in isolate P11B).
FT   VARIANT     590    590       T -> I (in isolate P100E).
FT   VARIANT     642    642       R -> M (in isolate A162).
FT   VARIANT     681    681       T -> R (in isolate A162).
FT   VARIANT     700    700       L -> I (in isolates RW Stock, P11A, P11B
FT                                 and P100E).
FT   VARIANT     711    711       D -> N (in isolate LRI 281).
FT   VARIANT     714    714       K -> N (in isolates RW Stock, P11A, P11B
FT                                 and P100E).
FT   VARIANT     765    765       V -> A (in isolate A162).
FT   VARIANT     775    775       A -> S (in isolate A162).

  Query Match             54.5%;  Score 54;  DB 1;  Length 1173;
  Best Local Similarity   85.7%;  Pred. No. 9.9;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 KWPWWPW 11
              ||||| |
Db       1113 KWPWWVW 1119


RESULT 3
Y013_NPVAC
ID   Y013_NPVAC     STANDARD;      PRT;   327 AA.
AC   P41423;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Hypothetical 38.7 kDa protein in PK1-LEF1 intergenic region.
OS   Autographa californica nuclear polyhedrosis virus (AcMNPV).
OC   Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC   Nucleopolyhedrovirus.
OX   NCBI_TaxID=46015;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C6;
RX   MEDLINE=94303173; PubMed=8030224;
RA   Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
RT   "The complete DNA sequence of Autographa californica nuclear
RT   polyhedrosis virus.";
RL   Virology 202:586-605(1994).
RN   [2]
RP   SEQUENCE OF 1-209 FROM N.A.
RC   STRAIN=L1;
RX   MEDLINE=93267802; PubMed=8497062;
RA   Passarelli A.L., Miller L.K.;
RT   "Identification and characterization of lef-1, a baculovirus gene
RT   involved in late and very late gene expression.";
RL   J. Virol. 67:3481-3488(1993).
CC   -!- SIMILARITY: TO CORRESPONDING ORF IN OPMNPV.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L22858; AAA66643.1; -.
DR   EMBL; L09723; AAA46706.1; -.
DR   PIR; E72851; E72851.
KW   Hypothetical protein.
SQ   SEQUENCE   327 AA;  38660 MW;  4D494C1D62285171 CRC64;

  Query Match             53.5%;  Score 53;  DB 1;  Length 327;
  Best Local Similarity   54.5%;  Pred. No. 4.2;
  Matches    6;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPW 11
              :| | | || |
Db          1 MLSWLWNWWMW 11


RESULT 4
BMR2_HUMAN
ID   BMR2_HUMAN     STANDARD;      PRT;  1038 AA.
AC   Q13873; Q16569;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Bone morphogenetic protein receptor type II precursor (EC 2.7.1.37)
DE   (BMP type II receptor) (BMPR-II).
```

```
GN   BMPR2 OR PPH1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Substantia nigra;
RX   MEDLINE=95372334; PubMed=7644468;
RA   Rosenzweig B.L., Imamura T., Okadome T., Cox G.N., Yamashita H.,
RA   ten Dijke P., Heldin C., Miyazono K.;
RT   "Cloning and characterization of a human type II receptor for bone
RT   morphogenetic proteins.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:7632-7636(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skin fibroblast;
RX   MEDLINE=95403457; PubMed=7673243;
RA   Nohno T., Ishikawa T., Saito T., Hosokawa K., Noji S., Wosing D.H.,
RA   Rosenbaum J.S.;
RT   "Identification of a human type II receptor for bone morphogenetic
RT   protein-4 that forms differential heteromeric complexes with bone
RT   morphogenetic protein type I receptors.";
RL   J. Biol. Chem. 270:22522-22526(1995).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95197572; PubMed=7890683;
RA   Kawabata M., Chytil A., Moses H.L.;
RT   "Cloning of a novel type II serine/threonine kinase receptor through
RT   interaction with the type I transforming growth factor-beta
RT   receptor.";
RL   J. Biol. Chem. 270:5625-5630(1995).
RN   [4]
RP   VARIANTS PPH1 GLN-491 AND TRP-491.
RX   MEDLINE=20395844; PubMed=10903931;
RA   Deng Z., Morse J.H., Slager S.L., Cuervo N., Moore K.J., Venetos G.,
RA   Kalachikov S., Cayanis E., Fischer S.G., Barst R.J., Hodge S.E.,
RA   Knowles J.A.;
RT   "Familial primary pulmonary hypertension (gene PPH1) is caused by
RT   mutations in the bone morphogenetic protein receptor-II gene.";
RL   Am. J. Hum. Genet. 67:737-744(2000).
RN   [5]
RP   VARIANTS PPH1 TYR-60; TYR-117 AND ARG-483.
RX   MEDLINE=20473811; PubMed=11015450;
RA   Thomson J.R., Machado R.D., Pauciulo M.W., Morgan N.V., Humbert M.,
RA   Elliott G.C., Ward K., Yacoub M., Mikhail G., Rogers P., Newman J.H.,
RA   Wheeler L., Higenbottam T., Gibbs J.S.R., Egan J., Crozier A.,
RA   Peacock A., Allcock R., Corris P., Loyd J.E., Trembath R.C.,
RA   Nichols W.C.;
RT   "Sporadic primary pulmonary hypertension is associated with germline
RT   mutations of the gene encoding BMPR-II, a receptor member of the
RT   TGF-beta family.";
RL   J. Med. Genet. 37:741-745(2000).
RN   [6]
RP   VARIANTS PPH1 TRP-118; TYR-347 AND GLY-485.
RX   MEDLINE=20428187; PubMed=10973254;
RA   Lane K.B., Machado R.D., Pauciulo M.W., Thomson J.R.,
RA   Phillips J.A. III, Loyd J.E., Nichols W.C., Trembath R.C., Aldred M.,
RA   Brannon C.A., Conneally P.M., Foroud T., Fretwell N., Gaddipati R.,
RA   Koller D., Loyd E.J., Morgan N.V., Newman J.H., Prince M.A.,
RA   Vilarino Gueell C., Wheeler L.;
RT   "Heterozygous germline mutations in BMPR2, encoding a TGF-beta
RT   receptor, cause familial primary pulmonary hypertension.";
RL   Nat. Genet. 26:81-84(2000).
RN   [7]
RP   VARIANTS PPH1 ARG-123; SER-123; ARG-420 AND THR-512, VARIANT ASP-224,
RP   AND CHARACTERIZATION OF VARIANT PPH1 GLY-485.
RX   MEDLINE=21063176; PubMed=11115378;
RA   Machado R.D., Pauciulo M.W., Thomson J.R., Lane K.B., Morgan N.V.,
RA   Wheeler L., Phillips J.A. III, Newman J.H., Williams D., Galie N.,
RA   Manes A., McNeil K., Yacoub M., Mikhail G., Rogers P., Corris P.,
RA   Humbert M., Donnai D., Martensson G., Tranebjaerg L., Loyd J.E.,
RA   Trembath R.C., Nichols W.C.;
RT   "BMPR2 haploinsufficiency as the inherited molecular mechanism for
RT   primary pulmonary hypertension.";
RL   Am. J. Hum. Genet. 68:92-102(2001).
CC   -!- FUNCTION: BINDS TO BMP-7, BMP-2 AND, LESS EFFICIENTLY, BMP-4.
CC       BINDING IS WEAK BUT ENHANCED BY THE PRESENCE OF TYPE I RECEPTORS
CC       FOR BMPS.
CC   -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC   -!- SUBUNIT: HETERODIMERIZE WITH TYPE-I RECEPTORS.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN HEART AND LIVER.
CC   -!- DISEASE: Defects in BMPR2 are the cause of primary pulmonary
CC       hypertension (PPH1) [MIM:178600]; a rare autosomal dominant
CC       disorder characterized by plexiform lesions of proliferating
CC       endothelial cells in pulmonary arterioles. The lesions lead to
CC       elevated pulmonary arterial pression, right ventricular failure,
CC       and death. The disease can occur from infancy throughout life and
CC       it has a mean age at onset of 36 years. Penetrance is reduced.
CC       Although familial PPH1 is rare, cases secondary to known
CC       etiologies are more common and include those associated with the
CC       appetite-suppressant drugs.
CC   -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC       TGFB RECEPTOR SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z48923; CAA88759.1; -.
DR   EMBL; D50516; BAA09094.1; -.
DR   EMBL; U20165; AAC50105.1; -.
DR   PIR; I38935; I38935.
DR   Genew; HGNC:1078; BMPR2.
DR   MIM; 600799; -.
DR   MIM; 178600; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0005515; F:protein binding activity; TAS.
DR   GO; GO:0007178; P:transmembrane receptor protein serine/threo. . .; TAS.
DR   InterPro; IPR000472; Activin_rec.
DR   InterPro; IPR000719; Prot_kinase.
DR   InterPro; IPR002290; Ser_thr_pkinase.
DR   Pfam; PF01064; Activin_recp; 1.
DR   Pfam; PF00069; pkinase; 1.
DR   ProDom; PD000001; Prot_kinase; 1.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW   Receptor; Transferase; Serine/threonine-protein kinase; ATP-binding;
KW   Transmembrane; Glycoprotein; Signal; Polymorphism; Disease mutation.
FT   SIGNAL       1     26       POTENTIAL.
FT   CHAIN       27   1038       BONE MORPHOGENETIC PROTEIN RECEPTOR TYPE
FT                               II.
FT   DOMAIN      27    150       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   151    171       POTENTIAL.
FT   DOMAIN     172   1038       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN     203    504       PROTEIN KINASE.
FT   NP_BIND    209    217       ATP (BY SIMILARITY).
FT   BINDING    230    230       ATP (BY SIMILARITY).
FT   ACT_SITE   333    333       BY SIMILARITY.
FT   DOMAIN     547    550       POLY-SER.
FT   DOMAIN     610    618       POLY-THR.
FT   DOMAIN     901    908       POLY-ASN.
FT   CARBOHYD    55     55       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   110    110       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   126    126       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT     60     60       C -> Y (in PPH1).
FT                               /FTId=VAR_013670.
FT   VARIANT    117    117       C -> Y (in PPH1).
FT                               /FTId=VAR_013671.
FT   VARIANT    118    118       C -> W (in PPH1).
```

```
FT                                   /FTId=VAR_013672.
FT   VARIANT     123    123          C -> R (in PPH1).
FT                                   /FTId=VAR_013673.
FT   VARIANT     123    123          C -> S (in PPH1).
FT                                   /FTId=VAR_013674.
FT   VARIANT     224    224          E -> D.
FT                                   /FTId=VAR_013675.
FT   VARIANT     347    347          C -> Y (in PPH1).
FT                                   /FTId=VAR_013676.
FT   VARIANT     420    420          C -> R (in PPH1).
FT                                   /FTId=VAR_013677.
FT   VARIANT     483    483          C -> R (in PPH1; sporadic).
FT                                   /FTId=VAR_013678.
FT   VARIANT     485    485          D -> G (in PPH1; complete loss of
FT                                   function).
FT                                   /FTId=VAR_013679.
FT   VARIANT     491    491          R -> Q (in PPH1; sporadic).
FT                                   /FTId=VAR_013680.
FT   VARIANT     491    491          R -> W (in PPH1).
FT                                   /FTId=VAR_013681.
FT   VARIANT     512    512          K -> T (in PPH1).
FT                                   /FTId=VAR_013682.
FT   VARIANT     519    519          N -> K (in PPH1).
FT                                   /FTId=VAR_013683.
FT   CONFLICT    828    828          G -> R (IN REF. 1).
SQ   SEQUENCE   1038 AA;  115201 MW;  1389923CE574B913 CRC64;

  Query Match             51.5%;  Score 51;  DB 1;  Length 1038;
  Best Local Similarity   66.7%;  Pred. No. 21;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 PWKWPWWPW 11
              ||: || ||
Db          8 PWRVPWLPW 16


RESULT 5
ATP8_CORCN
ID   ATP8_CORCN     STANDARD;      PRT;    55 AA.
AC   Q9TBI6;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A6L).
GN   MTATP8 OR ATP8.
OS   Corythaixoides concolor (Grey go-away-bird).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Musophagiformes; Musophagidae;
OC   Corythaixoides.
OX   NCBI_TaxID=103956;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99416451; PubMed=10486983;
RA   Hughes J.M., Baker A.J.;
RT   "Phylogenetic relationships of the enigmatic hoatzin (Opisthocomus
RT   hoazin) resolved using mitochondrial and nuclear gene sequences.";
RL   Mol. Biol. Evol. 16:1300-1307(1999).
CC   -!- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
CC       (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) = ADP + phosphate +
CC       H(+)(Out).
CC   -!- SUBCELLULAR LOCATION: Membrane-bound.
CC   -!- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF168039; AAD56467.1; -.
DR   InterPro; IPR001421; ATPase8_mit.
DR   Pfam; PF00895; ATP-synt_8; 1.
KW   Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
SQ   SEQUENCE   55 AA;  6485 MW;  973552DB00E918AD CRC64;

  Query Match             50.5%;  Score 50;  DB 1;  Length 55;
  Best Local Similarity   85.7%;  Pred. No. 2;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 LPWKWPW 8
              ||| |||
Db         48 LPWTWPW 54


RESULT 6
ADRO_BOVIN
ID   ADRO_BOVIN     STANDARD;      PRT;   492 AA.
AC   P08165; Q95KN8;
DT   01-AUG-1988 (Rel. 08, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   NADPH:adrenodoxin oxidoreductase, mitochondrial precursor
DE   (EC 1.18.1.2) (Adrenodoxin reductase) (AR) (Ferredoxin-NADP(+)
DE   reductase).
GN   FDXR OR ADXR.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX   MEDLINE=94177140; PubMed=8130767;
RA   Takata Y., Sagara Y., Kono A., Sekimizu K., Horiuchi T.;
RT   "Gene structure of bovine adrenodoxin reductase.";
RL   Biol. Pharm. Bull. 16:1200-1206(1993).
RN   [2]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=88198050; PubMed=3448086;
RA   Sagara Y., Takata Y., Miyata T., Hara T., Horiuchi T.;
RT   "Cloning and sequence analysis of adrenodoxin reductase cDNA from
RT   bovine adrenal cortex.";
RL   J. Biochem. 102:1333-1336(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87270696; PubMed=3038094;
RA   Nonaka Y., Murakami H., Yabusaki Y., Kuramitsu S., Kagamiyama H.,
RA   Yamano T., Okamoto M.;
RT   "Molecular cloning and sequence analysis of full-length cDNA for mRNA
RT   of adrenodoxin oxidoreductase from bovine adrenal cortex.";
RL   Biochem. Biophys. Res. Commun. 145:1239-1247(1987).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Adrenal cortex;
RX   MEDLINE=89170752; PubMed=2924777;
RA   Hanukoglu I., Gutfinger T.;
RT   "cDNA sequence of adrenodoxin reductase. Identification of NADP-
RT   binding sites in oxidoreductases.";
RL   Eur. J. Biochem. 180:479-484(1989).
RN   [5]
RP   SEQUENCE OF N-TERMINUS, AND PARTIAL SEQUENCE.
RC   TISSUE=Adrenal cortex;
RX   MEDLINE=88082777; PubMed=3691502;
RA   Hanukoglu I., Gutfinger T., Haniu M., Shively J.E.;
RT   "Isolation of a cDNA for adrenodoxin reductase (ferredoxin-NADP+
RT   reductase). Implications for mitochondrial cytochrome P-450 systems.";
RL   Eur. J. Biochem. 169:449-455(1987).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 33-492.
RC   TISSUE=Adrenal gland;
RX   MEDLINE=99299392; PubMed=10369776;
RA   Ziegler G.A., Vonrhein C., Hanukoglu I., Schulz G.E.;
```

```
RT   "The structure of adrenodoxin reductase of mitochondrial P450 systems:
RT   electron transfer for steroid biosynthesis.";
RL   J. Mol. Biol. 289:981-990(1999).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).
RX   MEDLINE=20455764; PubMed=10998235;
RA   Ziegler G.A., Schulz G.E.;
RT   "Crystal structures of adrenodoxin reductase in complex with NADP+ and
RT   NADPH suggesting a mechanism for the electron transfer of an enzyme
RT   family.";
RL   Biochemistry 39:10986-10995(2000).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF COMPLEX WITH ADRENODOXIN.
RX   MEDLINE=21264735; PubMed=11053423;
RA   Mueller J.J., Lapko A., Bourenkov G., Ruckpaul K., Heinemann U.;
RT   "Adrenodoxin reductase-adrenodoxin complex structure suggests electron
RT   transfer path in steroid biosynthesis.";
RL   J. Biol. Chem. 276:2786-2789(2001).
CC   -!- FUNCTION: SERVES AS THE FIRST ELECTRON TRANSFER PROTEIN IN ALL THE
CC       MITOCHONDRIAL P450 SYSTEMS. INCLUDING CHOLESTEROL SIDE CHAIN
CC       CLEAVAGE IN ALL STEROIDOGENIC TISSUES, STEROID 11-BETA
CC       HYDROXYLATION IN THE ADRENAL CORTEX, 25-OH-VITAMIN D3-24
CC       HYDROXYLATION IN THE KIDNEY, AND STEROL C-27 HYDROXYLATION IN THE
CC       LIVER.
CC   -!- CATALYTIC ACTIVITY: Reduced adrenodoxin + NADP(+) = oxidized
CC       adrenodoxin + NADPH.
CC   -!- COFACTOR: FAD.
CC   -!- PATHWAY: CHOLESTEROL SIDE-CHAIN-CLEAVAGE SYSTEM.
CC   -!- SUBUNIT: Monomer.
CC   -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=Short;
CC         IsoId=P08165-1; Sequence=Displayed;
CC       Name=Long;
CC         IsoId=P08165-2; Sequence=VSP_003415;
CC         Note=Represents 10-20% of all adrenodoxin reductase mRNAs and
CC         seems to be inactive;
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D83475; BAA11921.1; -.
DR   EMBL; D83472; BAA11921.1; JOINED.
DR   EMBL; D83473; BAA11921.1; JOINED.
DR   EMBL; D83474; BAA11921.1; JOINED.
DR   EMBL; M17029; AAA30362.1; -.
DR   EMBL; D00211; BAA00150.1; -.
DR   EMBL; X13736; CAA32002.1; -.
DR   PIR; JT0751; JT0751.
DR   PDB; 1CJC; 12-APR-99.
DR   PDB; 1E1L; 24-SEP-00.
DR   PDB; 1E1K; 24-SEP-00.
DR   PDB; 1E1M; 24-SEP-00.
DR   PDB; 1E1N; 24-SEP-00.
DR   PDB; 1E6E; 09-AUG-01.
DR   InterPro; IPR000759; Adrndx_reductase.
DR   PRINTS; PR00419; ADXRDTASE.
KW   Electron transport; Oxidoreductase; Flavoprotein; NADP; FAD;
KW   Mitochondrion; Transit peptide; Alternative splicing; 3D-structure.
FT   TRANSIT       1     32       MITOCHONDRION.
FT   CHAIN        33    492       NADPH:ADRENODOXIN OXIDOREDUCTASE.
FT   VARSPLIC    204    204       E -> EVLLLCQ (in isoform Long).
FT                                /FTId=VSP_003415.
FT   CONFLICT     77     77       G -> R (IN REF. 3).
FT   CONFLICT     81     94       FGVAPDHPEVKNVI -> VWLALTTPRSRMLL (IN REF.
FT                                3).
FT   CONFLICT    124    128       QDAYH -> RVYRLT (IN REF. 3).
FT   CONFLICT    268    268       K -> R (IN REF. 3).
FT   CONFLICT    317    318       PS -> RL (IN REF. 3).
FT   CONFLICT    323    333       RAAGIRLAVTR -> ARRSAWQSPE (IN REF. 3).
FT   CONFLICT    341    352       TRAVPTGDVEDL -> HPGSAHWGCGGP (IN REF. 3).
FT   STRAND       40     44
FT   HELIX        48     60
FT   STRAND       65     69
FT   TURN         77     77
FT   HELIX        78     81
FT   TURN         82     82
FT   TURN         85     86
FT   HELIX        88     92
FT   HELIX        93    101
FT   TURN        102    102
FT   TURN        104    105
FT   STRAND      106    110
FT   STRAND      114    114
FT   TURN        115    117
FT   STRAND      118    118
FT   HELIX       120    126
FT   STRAND      129    132
FT   STRAND      138    139
FT   TURN        145    148
FT   TURN        150    151
FT   STRAND      152    154
FT   HELIX       155    162
FT   TURN        163    164
FT   HELIX       166    168
FT   TURN        169    170
FT   TURN        175    176
FT   STRAND      179    183
FT   HELIX       187    197
FT   HELIX       200    203
FT   TURN        204    205
FT   HELIX       210    217
FT   TURN        218    218
FT   STRAND      223    227
FT   HELIX       232    234
FT   HELIX       239    246
FT   TURN        247    247
FT   TURN        249    250
FT   STRAND      251    254
FT   HELIX       257    260
FT   TURN        261    262
FT   HELIX       263    266
FT   TURN        267    269
FT   HELIX       272    286
FT   HELIX       291    298
FT   TURN        299    299
FT   STRAND      302    307
FT   STRAND      310    317
FT   TURN        319    320
FT   STRAND      324    335
FT   HELIX       338    340
FT   STRAND      342    353
FT   STRAND      356    359
FT   STRAND      363    364
FT   TURN        370    371
FT   STRAND      375    375
FT   TURN        376    379
FT   STRAND      380    380
FT   STRAND      383    384
FT   TURN        385    386
FT   STRAND      387    388
FT   TURN        389    390
FT   TURN        392    393
FT   STRAND      394    396
FT   TURN        398    398
FT   HELIX       399    402
FT   TURN        404    405
FT   HELIX       408    428

  Query Match          50.5%;  Score 50;  DB 1;  Length 492;
```

```
  Best Local Similarity   66.7%;  Pred. No. 14;
  Matches    6;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 PWKWPWWPW 11
              |  | ||||
Db          3 PRCWRWWPW 11


RESULT 7
ATP8_PELSU
ID   ATP8_PELSU     STANDARD;      PRT;    55 AA.
AC   O79674;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A6L).
GN   MTATP8 OR ATP8.
OS   Pelomedusa subrufa (African side-necked turtle).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Testudines; Pleurodira; Pelomedusidae; Pelomedusa.
OX   NCBI_TaxID=44522;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zardoya R.;
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
CC       (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) = ADP + phosphate +
CC       H(+)(Out).
CC   -!- SUBCELLULAR LOCATION: Membrane-bound.
CC   -!- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF039066; AAD05054.1; -.
DR   PIR; T11105; T11105.
DR   InterPro; IPR001421; ATPase8_mit.
DR   Pfam; PF00895; ATP-synt_8; 1.
KW   Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
FT   TRANSMEM      4     24       POTENTIAL.
SQ   SEQUENCE   55 AA;  6536 MW;  D8D4BC8F8651A001 CRC64;

  Query Match             48.5%;  Score 48;  DB 1;  Length 55;
  Best Local Similarity   71.4%;  Pred. No. 3.5;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 LPWKWPW 8
              :|| |||
Db         48 MPWTWPW 54


RESULT 8
ELO1_HUMAN
ID   ELO1_HUMAN     STANDARD;      PRT;    279 AA.
AC   Q9BW60; Q9NVD9; Q9Y396;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Elongation of very long chain fatty acids protein 1 (CGI-88).
GN   ELOVL1 OR SSC1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20272150; PubMed=10810093;
RA   Lai C.-H., Chou C.-Y., Ch'ang L.-Y., Liu C.-S., Lin W.-C.;
RT   "Identification of novel human genes evolutionarily conserved in
RT   Caenorhabditis elegans by comparative proteomics.";
RL   Genome Res. 10:703-713(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA   Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
RA   Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: Could be implicated in tissue-specific synthesis of very
CC       long chain fatty acids and sphingolipids. May catalyze one or both
CC       of the reduction reaction in fatty acid elongation, i.e.,
CC       conversion of beta-ketoacyl CoA to beta-hydroxyacyl CoA or
CC       reduction of trans-2-enoyl CoA to the saturated acyl CoA
CC       derivative (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC       reticulum (Potential).
CC   -!- SIMILARITY: BELONGS TO THE ELO FAMILY.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC       frameshift in position 189.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF151846; AAD34083.1; ALT_FRAME.
DR   EMBL; AK001653; BAA91813.1; -.
DR   EMBL; BC000618; AAH00618.1; -.
DR   Genew; HGNC:14418; ELOVL1.
DR   InterPro; IPR002076; GNS1_SUR4.
DR   Pfam; PF01151; ELO; 1.
DR   PROSITE; PS01188; ELO; 1.
KW   Fatty acid biosynthesis; Transmembrane; Endoplasmic reticulum.
FT   TRANSMEM     23     43       POTENTIAL.
FT   TRANSMEM     61     81       POTENTIAL.
FT   TRANSMEM    176    196       POTENTIAL.
FT   TRANSMEM    201    221       POTENTIAL.
FT   TRANSMEM    231    251       POTENTIAL.
FT   SITE         275    277       ENDOPLASMIC RETICULUM RETRIEVAL MOTIF
FT                                 (POTENTIAL).
FT   CONFLICT     68     68       S -> P (IN REF. 2).
```

```
SQ   SEQUENCE   279 AA;  32663 MW;  B168EE4C7EAF92A6 CRC64;

  Query Match             48.0%;  Score 47.5;  DB 1;  Length 279;
  Best Local Similarity   66.7%;  Pred. No. 17;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          1 ILPWKWPWW 9
              :||| | ||
Db        147 VLPWSW-WW 154


RESULT 9
ELO1_MOUSE
ID   ELO1_MOUSE     STANDARD;      PRT;   279 AA.
AC   Q9JLJ5; Q9D1B2;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Elongation of very long chain fatty acids protein 1.
GN   ELOVL1 OR SSC1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Liver;
RX   MEDLINE=20253178; PubMed=10791983;
RA   Tvrdik P., Westerberg R., Silve S., Asadi A., Jakobsson A., Cannon B.,
RA   Loison G., Jacobsson A.;
RT   "Role of a new mammalian gene family in the biosynthesis of very long
RT   chain fatty acids and sphingolipids.";
RL   J. Cell Biol. 149:707-718(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Breast tumor;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   SEQUENCE OF 78-279 FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Embryo;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
CC   -!- FUNCTION: Could be implicated in tissue-specific synthesis of very
CC       long chain fatty acids and sphingolipids. May catalyze one or both
CC       of the reduction reaction in fatty acid elongation, i.e.,
CC       conversion of beta-ketoacyl CoA to beta-hydroxyacyl CoA or
CC       reduction of trans-2-enoyl CoA to the saturated acyl CoA
CC       derivative.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC       reticulum (Potential).
CC   -!- TISSUE SPECIFICITY: Expressed in a broad variety of tissues.
CC       Highly expressed in stomach, lung, kidney, skin and intestine.
CC       Moderately expressed in white adipose tissue, liver, spleen,
CC       brain, brown adipose tissue, heart and muscle. Weakly expressed in
CC       testis.
CC   -!- SIMILARITY: BELONGS TO THE ELO FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF170907; AAF72572.1; -.
DR   EMBL; BC006735; AAH06735.1; -.
DR   EMBL; AK003743; BAB22975.1; -.
DR   MGD; MGI:1858959; Elovl1.
DR   InterPro; IPR002076; GNS1_SUR4.
DR   Pfam; PF01151; ELO; 1.
DR   PROSITE; PS01188; ELO; 1.
KW   Fatty acid biosynthesis; Transmembrane; Endoplasmic reticulum.
FT   TRANSMEM     23     43       POTENTIAL.
FT   TRANSMEM     61     81       POTENTIAL.
FT   TRANSMEM    176    196       POTENTIAL.
FT   TRANSMEM    203    223       POTENTIAL.
FT   TRANSMEM    231    251       POTENTIAL.
FT   SITE        275    277       ENDOPLASMIC RETICULUM RETRIEVAL MOTIF
FT                                (POTENTIAL).
FT   CONFLICT     78     79       YE -> MR (IN REF. 3).
SQ   SEQUENCE   279 AA;  32678 MW;  CA5A1CF55FDB2F76 CRC64;

  Query Match             48.0%;  Score 47.5;  DB 1;  Length 279;
  Best Local Similarity   66.7%;  Pred. No. 17;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          1 ILPWKWPWW 9
              :||| | ||
Db        147 VLPWSW-WW 154


RESULT 10
ATP8_GADMO
ID   ATP8_GADMO     STANDARD;      PRT;    55 AA.
AC   P15996;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A6L).
GN   MTATP8 OR ATP8.
OS   Gadus morhua (Atlantic cod).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
OX   NCBI_TaxID=8049;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Norwegian coastal 1; TISSUE=Liver;
RX   MEDLINE=90174958; PubMed=2308841;
RA   Johansen S., Guddal P.H., Johansen T.;
RT   "Organization of the mitochondrial genome of Atlantic cod, Gadus
RT   morhua.";
RL   Nucleic Acids Res. 18:411-419(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Norwegian coastal 1;
RX   MEDLINE=96414925; PubMed=8817926;
RA   Johansen S., Bakke I.;
RT   "The complete mitochondrial DNA sequence of Atlantic cod (Gadus
RT   morhua): relevance to taxonomic studies among codfishes.";
RL   Mol. Mar. Biol. Biotechnol. 5:203-214(1996).
CC   -!- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
CC       (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) = ADP + phosphate +
CC       H(+)(Out).
CC   -!- SUBCELLULAR LOCATION: Membrane-bound.
CC   -!- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X17659; CAA35655.1; -.
DR   EMBL; X99772; CAA68110.1; -.
DR   PIR; S08424; S08424.
DR   InterPro; IPR001421; ATPase8_mit.
DR   Pfam; PF00895; ATP-synt_8; 1.
KW   Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
FT   TRANSMEM      4     24       POTENTIAL.
SQ   SEQUENCE   55 AA;  6481 MW;  E85C81E63DB48B15 CRC64;

  Query Match             47.5%;  Score 47;  DB 1;  Length 55;
  Best Local Similarity   83.3%;  Pred. No. 4.7;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 PWKWPW 8
              || |||
Db         49 PWNWPW 54


RESULT 11
ATP8_ONCMY
ID   ATP8_ONCMY     STANDARD;      PRT;    55 AA.
AC   P48179;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A6L).
GN   MTATP8 OR ATP8.
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri), and
OS   Salmo salar (Atlantic salmon).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8022, 8030;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=O.mykiss; TISSUE=Liver;
RX   MEDLINE=96139027; PubMed=8587139;
RA   Zardoya R., Garrido-Pertierra A., Bautista J.M.;
RT   "The complete nucleotide sequence of the mitochondrial DNA genome of
RT   the rainbow trout, Oncorhynchus mykiss.";
RL   J. Mol. Evol. 41:942-951(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   SPECIES=S.salar; TISSUE=Liver;
RX   MEDLINE=20018174; PubMed=10548724;
RA   Hurst C.D., Bartlett S.E., Davidson W.S., Bruce I.J.;
RT   "The complete mitochondrial DNA sequence of the Atlantic salmon, Salmo
RT   salar.";
RL   Gene 239:237-242(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   SPECIES=S.salar;
RA   Arnason U., Johnsson E., Rasmussen A.S.;
RT   "The complete mitochondrial genome sequence of a teleost, Salmo salar,
RT   and comparisons with other salmoniformes.";
RL   Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
CC       (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) = ADP + phosphate +
CC       H(+)(Out).
CC   -!- SUBCELLULAR LOCATION: Membrane-bound.
CC   -!- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L29771; AAB03351.1; -.
DR   EMBL; U12143; AAD04737.1; -.
DR   EMBL; AF133701; AAF61382.1; -.
DR   PIR; T09861; T09861.
DR   PIR; T09951; T09951.
DR   InterPro; IPR001421; ATPase8_mit.
DR   Pfam; PF00895; ATP-synt_8; 1.
KW   Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
FT   TRANSMEM      4     24       POTENTIAL.
SQ   SEQUENCE   55 AA;  6413 MW;  D02920C3E346925F CRC64;

  Query Match             47.5%;  Score 47;  DB 1;  Length 55;
  Best Local Similarity   83.3%;  Pred. No. 4.7;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 PWKWPW 8
              || |||
Db         49 PWNWPW 54


RESULT 12
ATP8_PRODO
ID   ATP8_PRODO     STANDARD;      PRT;    55 AA.
AC   Q35416;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A6L).
GN   MTATP8 OR ATP8.
OS   Protopterus dolloi (Lungfish).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Dipnoi; Lepidosireniformes; Protopteridae; Protopterus.
OX   NCBI_TaxID=27779;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Egg;
RX   MEDLINE=96271539; PubMed=8846902;
RA   Zardoya R., Meyer A.;
RT   "The complete nucleotide sequence of the mitochondrial genome of the
RT   lungfish (Protopterus dolloi) supports its phylogenetic position as a
RT   close relative of land vertebrates.";
RL   Genetics 142:1249-1263(1996).
```

```
CC   -!- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
CC       (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) = ADP + phosphate +
CC       H(+)(Out).
CC   -!- SUBCELLULAR LOCATION: Membrane-bound.
CC   -!- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L42813; AAC38025.1; -.
DR   PIR; S68132; S68132.
DR   InterPro; IPR001421; ATPase8_mit.
DR   Pfam; PF00895; ATP-synt_8; 1.
KW   Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
SQ   SEQUENCE   55 AA;  6523 MW;  95343043B5B2DC53 CRC64;

  Query Match             47.5%;  Score 47;  DB 1;  Length 55;
  Best Local Similarity   83.3%;  Pred. No. 4.7;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 PWKWPW 8
              || |||
Db         49 PWNWPW 54


RESULT 13
ATP8_SALAL
ID   ATP8_SALAL     STANDARD;      PRT;    55 AA.
AC   Q9XN27;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A6L).
GN   MTATP8 OR ATP8 OR ATPASE8.
OS   Salvelinus alpinus (Arctic char).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
OX   NCBI_TaxID=8036;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Doiron S., Blier P.U., Bernatchez L.;
RT   "A comparative analysis of complete sequence of mitochondrial genome
RT   between brook char (Salvelinus fontinalis) and arctic char (S.
RT   alpinus).";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
CC       (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) = ADP + phosphate +
CC       H(+)(Out).
CC   -!- SUBCELLULAR LOCATION: Membrane-bound.
CC   -!- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF154851; AAD41389.1; -.
DR   InterPro; IPR001421; ATPase8_mit.
DR   Pfam; PF00895; ATP-synt_8; 1.
KW   Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
SQ   SEQUENCE   55 AA;  6455 MW;  71E430C2E346924A CRC64;

  Query Match             47.5%;  Score 47;  DB 1;  Length 55;
  Best Local Similarity   83.3%;  Pred. No. 4.7;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 PWKWPW 8
              || |||
Db         49 PWNWPW 54


RESULT 14
ATP8_SALFO
ID   ATP8_SALFO     STANDARD;      PRT;    55 AA.
AC   Q9XN35;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A6L).
GN   MTATP8 OR ATP8 OR ATPASE8.
OS   Salvelinus fontinalis (Brook trout) (Brook char).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
OX   NCBI_TaxID=8038;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Doiron S., Blier P.U., Bernatchez L.;
RT   "A comparative analysis of complete sequence of mitochondrial genome
RT   between brook char (Salvelinus fontinalis) and arctic char (S.
RT   alpinus).";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
CC       (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) = ADP + phosphate +
CC       H(+)(Out).
CC   -!- SUBCELLULAR LOCATION: Membrane-bound.
CC   -!- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF154850; AAD41376.1; -.
DR   InterPro; IPR001421; ATPase8_mit.
DR   Pfam; PF00895; ATP-synt_8; 1.
KW   Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
SQ   SEQUENCE   55 AA;  6443 MW;  D02930C2E346925F CRC64;

  Query Match             47.5%;  Score 47;  DB 1;  Length 55;
  Best Local Similarity   83.3%;  Pred. No. 4.7;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 PWKWPW 8
              || |||
Db         49 PWNWPW 54


RESULT 15
ATP8_SCYCA
ID   ATP8_SCYCA     STANDARD;      PRT;    55 AA.
AC   O79405;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   ATP synthase protein 8 (EC 3.6.3.14) (ATPase subunit 8) (A6L).
GN   MTATP8 OR ATP8 OR ATPASE8.
OS   Scyliorhinus canicula (Spotted dogfish) (Spotted catshark).
OG   Mitochondrion.
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC   Scyliorhinidae; Scyliorhinus.
OX   NCBI_TaxID=7830;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Muscle;
RX   MEDLINE=98393590; PubMed=9725850;
RA   Delarbre C., Spruyt N., Delmarre C., Gallut C., Barriel V.,
RA   Janvier P., Laudet V., Gachelin G.;
RT   "The complete nucleotide sequence of the mitochondrial DNA of the
RT   dogfish, Scyliorhinus canicula.";
RL   Genetics 150:331-344(1998).
CC   -!- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
CC       (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.
CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) = ADP + phosphate +
CC       H(+)(Out).
CC   -!- SUBCELLULAR LOCATION: Membrane-bound.
CC   -!- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y16067; CAA76023.1; -.
DR   PIR; T11304; T11304.
DR   InterPro; IPR001421; ATPase8_mit.
DR   Pfam; PF00895; ATP-synt_8; 1.
KW   Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
FT   TRANSMEM      4     24       POTENTIAL.
SQ   SEQUENCE   55 AA;  6607 MW;  075956C2A3DF05B9 CRC64;

  Query Match             47.5%;  Score 47;  DB 1;  Length 55;
  Best Local Similarity   83.3%;  Pred. No. 4.7;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 PWKWPW 8
              || |||
Db         49 PWNWPW 54


Search completed: October  2, 2003, 10:02:01
Job time : 25 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 2, 2003, 09:59:17 ; Search time 105 Seconds
(without alignments)
31.949 Million cell updates/sec

Title: US-09-444-281-85
Perfect score: 99
Sequence: 1 ILPWKWPWWPWRR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	59	59.6	780	16	Q8PE93	Q8pe93 xanthomonas
2	57	57.6	723	12	Q9DUC4	Q9duc4 tt virus. o
3	55	55.6	49	12	Q9DT80	Q9dt80 tt virus. o
4	55	55.6	152	2	Q8RPF4	Q8rpf4 desulfitoba
5	55	55.6	208	16	Q8PBI7	Q8pbi7 xanthomonas
6	55	55.6	342	4	Q96BE4	Q96be4 homo sapien
7	55	55.6	748	12	Q9DT81	Q9dt81 tt virus. o
8	55	55.6	750	12	Q91D04	Q91d04 tt virus. o
9	54	54.5	1173	12	Q990M4	Q990m4 human coron
10	54	54.5	1173	12	Q990M1	Q990m1 human coron
11	54	54.5	1173	12	Q990M3	Q990m3 human coron
12	54	54.5	1173	12	Q990M2	Q990m2 human coron
13	54	54.5	1383	12	Q84712	Q84712 porcine epi
14	54	54.5	1383	12	Q91AV1	Q91av1 porcine epi
15	54	54.5	1383	12	Q8B482	Q8b482 porcine epi
16	54	54.5	1386	12	Q8QQ98	Q8qq98 porcine epi
17	53.5	54.0	299	4	Q9Y4N1	Q9y4n1 homo sapien
18	53	53.5	216	5	Q9W476	Q9w476 drosophila
19	53	53.5	257	17	Q8TWW9	Q8tww9 methanopyru
20	53	53.5	328	12	Q8B9N6	Q8b9n6 rachiplusia
21	53	53.5	331	12	O92380	O92380 bombyx mori
22	53	53.5	600	5	Q8IGB8	Q8igb8 drosophila
23	53	53.5	1245	3	Q9Y7V5	Q9y7v5 trichoderma
24	52.5	53.0	640	2	Q934J3	Q934j3 prevotella
25	52	52.5	102	16	Q8P4Z9	Q8p4z9 xanthomonas
26	52	52.5	105	16	Q8PPU5	Q8ppu5 xanthomonas
27	52	52.5	351	16	Q8DJH5	Q8djh5 synechococc
28	51	51.5	55	8	Q9B6T0	Q9b6t0 eudromia el
29	51	51.5	298	17	Q8ZU59	Q8zu59 pyrobaculum
30	51	51.5	530	4	Q13161	Q13161 homo sapien
31	51	51.5	689	16	Q8YY85	Q8yy85 anabaena sp
32	50.5	51.0	214	5	Q9N9T4	Q9n9t4 leishmania
33	50.5	51.0	970	12	Q9YW19	Q9yw19 melanoplus
34	50.5	51.0	988	12	Q91HP7	Q91hp7 oedaleus as
35	50	50.5	55	8	Q8SEB4	Q8seb4 elaenia fal
36	50	50.5	83	16	Q9WYF1	Q9wyf1 thermotoga
37	50	50.5	327	10	Q9AUN3	Q9aun3 oryza sativ
38	50	50.5	337	16	Q92VQ2	Q92vq2 rhizobium m
39	50	50.5	466	4	O75035	O75035 homo sapien
40	50	50.5	746	12	Q9JH31	Q9jh31 tt virus. o
41	49.5	50.0	157	5	Q9Y0E8	Q9y0e8 drosophila
42	49.5	50.0	198	10	Q8GZX7	Q8gzx7 oryza sativ
43	49.5	50.0	296	2	O68125	O68125 rhodobacter
44	49.5	50.0	310	11	Q8CB00	Q8cb00 mus musculu
45	49.5	50.0	806	10	Q9FGM0	Q9fgm0 arabidopsis

ALIGNMENTS

RESULT 1
Q8PE93
ID Q8PE93 PRELIMINARY; PRT; 780 AA.
AC Q8PE93;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein XCC0088.
GN XCC0088.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL; AE012102; AAM39407.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 780 AA; 85074 MW; 12867434D1852549 CRC64;

```
  Query Match             59.6%;  Score 59;  DB 16;  Length 780;
  Best Local Similarity   75.0%;  Pred. No. 13;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 WKWPWWPW 11
              | | ||||
Db        148 WPWHWWPW 155


RESULT 2
Q9DUC4
ID   Q9DUC4      PRELIMINARY;      PRT;   723 AA.
AC   Q9DUC4;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   ORF1.
OS   TT virus.
OC   Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX   NCBI_TaxID=68887;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Mf-TTV9;
RA   Okamoto H.;
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Mf-TTV9;
RX   MEDLINE=20534983; PubMed=11080484;
RA   Okamoto H., Nishizawa T., Tawara A., Peng Y., Takahashi M.,
RA   Kishimoto J., Tanaka T., Miyakawa Y., Mayumi M.;
RT   "Species-specific TT viruses in humans and nonhuman primates and their
RT   phylogenetic relatedness.";
RL   Virology 277:368-378(2000).
DR   EMBL; AB041959; BAB19313.1; -.
DR   InterPro; IPR001563; Serine_carbpept.
DR   InterPro; IPR004219; TTvirus_Unk.
DR   Pfam; PF02956; TT_ORF1; 1.
DR   PROSITE; PS00131; CARBOXYPEPT_SER_SER; 1.
SQ   SEQUENCE   723 AA;  85393 MW;  232D003098766344 CRC64;

  Query Match             57.6%;  Score 57;  DB 12;  Length 723;
  Best Local Similarity   100.0%;  Pred. No. 22;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 PWWPWRR 13
              |||||||
Db          2 PWWPWRR 8


RESULT 3
Q9DT80
ID   Q9DT80      PRELIMINARY;      PRT;    49 AA.
AC   Q9DT80;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   ORF1 (Fragment).
OS   TT virus.
OC   Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX   NCBI_TaxID=68887;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TYM9;
RX   MEDLINE=20568739; PubMed=11118348;
RA   Okamoto H., Nishizawa T., Tawara A., Takahashi M., Kishimoto J.,
RA   Sai T., Sugai Y.;
RT   "TT virus mRNAs detected in the bone marrow cells from an infected
RT   individual.";
RL   Biochem. Biophys. Res. Commun. 279:700-707(2000).
DR   EMBL; AB050449; BAB19930.1; -.
FT   NON_TER      49      49
SQ   SEQUENCE   49 AA;  7225 MW;  1DA6F8F1AB69AA43 CRC64;

  Query Match             55.6%;  Score 55;  DB 12;  Length 49;
  Best Local Similarity   44.4%;  Pred. No. 3.6;
  Matches    8;  Conservative    1;  Mismatches    3;  Indels    6;  Gaps    1;

Qy          2 LPWKWPW------WPWRR 13
              : | | |      |||||
Db          1 MAWTWWWQRRRRRWPWRR 18


RESULT 4
Q8RPF4
ID   Q8RPF4      PRELIMINARY;      PRT;   152 AA.
AC   Q8RPF4;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical 16.9 kDa protein.
OS   Desulfitobacterium hafniense.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Peptococcaceae;
OC   Desulfitobacterium.
OX   NCBI_TaxID=49338;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DCB-2;
RA   Davis J.K., Tiedje J.M.;
RT   "Sequence and transcriptional analysis of reductive dehalogenase genes
RT   of Desulfitobacterium.";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF403185; AAL87800.1; -.
DR   InterPro; IPR006311; Tat.
DR   TIGRFAMs; TIGR01409; TAT_signal_seq; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   152 AA;  16876 MW;  2F5A00F01E70A379 CRC64;

  Query Match             55.6%;  Score 55;  DB 2;  Length 152;
  Best Local Similarity   85.7%;  Pred. No. 9.7;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 PWKWPWW 9
              || ||||
Db        146 PWIWPWW 152


RESULT 5
Q8PBI7
ID   Q8PBI7      PRELIMINARY;      PRT;   208 AA.
AC   Q8PBI7;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein XCC1132.
GN   XCC1132.
OS   Xanthomonas campestris (pv. campestris).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xanthomonas.
OX   NCBI_TaxID=340;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 33913 / NCPPB 528;
RX   MEDLINE=22022145; PubMed=12024217;
RA   da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA   Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA   Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA   Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA   Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA   Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA   Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA   Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA   Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA   Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA   Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA   Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
```

```
RA   Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA   Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA   Setubal J.C., Kitajima J.P.;
RT   "Comparison of the genomes of two Xanthomonas pathogens with differing
RT   host specificities.";
RL   Nature 417:459-463(2002).
DR   EMBL; AE012212; AAM40431.1; -.
DR   InterPro; IPR003709; VanY.
DR   Pfam; PF02557; VanY; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   208 AA;  22940 MW;  10B180F6EAF7B014 CRC64;

  Query Match             55.6%;  Score 55;  DB 16;  Length 208;
  Best Local Similarity   75.0%;  Pred. No. 13;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 PWKWPWWP 10
              || | |||
Db        200 PWHWRWWP 207


RESULT 6
Q96BE4
ID   Q96BE4      PRELIMINARY;      PRT;   342 AA.
AC   Q96BE4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skin, and amelanotic;
RA   Strausberg R.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC015687; AAH15687.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   342 AA;  37741 MW;  3147596F8D7DF849 CRC64;

  Query Match             55.6%;  Score 55;  DB 4;  Length 342;
  Best Local Similarity   63.6%;  Pred. No. 19;
  Matches    7;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPW 11
              :||  || |||
Db        296 LLPGPWPGWPW 306


RESULT 7
Q9DT81
ID   Q9DT81      PRELIMINARY;      PRT;   748 AA.
AC   Q9DT81;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   ORF1.
OS   TT virus.
OC   Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX   NCBI_TaxID=68887;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TYM9;
RX   MEDLINE=20568739; PubMed=11118348;
RA   Okamoto H., Nishizawa T., Tawara A., Takahashi M., Kishimoto J.,
RA   Sai T., Sugai Y.;
RT   "TT virus mRNAs detected in the bone marrow cells from an infected
RT   individual.";
RL   Biochem. Biophys. Res. Commun. 279:700-707(2000).
DR   EMBL; AB050448; BAB19928.1; -.
DR   InterPro; IPR004219; TTvirus_Unk.
DR   Pfam; PF02956; TT_ORF1; 1.
SQ   SEQUENCE   748 AA;  88552 MW;  D65CCB2CAA5CE26F CRC64;

  Query Match             55.6%;  Score 55;  DB 12;  Length 748;
  Best Local Similarity   44.4%;  Pred. No. 38;
  Matches    8;  Conservative    1;  Mismatches    3;  Indels    6;  Gaps    1;

Qy          2 LPWKWPW------WPWRR 13
              : | | |      |||||
Db          1 MAWTWWWQRRRRRWPWRR 18


RESULT 8
Q91D04
ID   Q91D04      PRELIMINARY;      PRT;   750 AA.
AC   Q91D04;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   ORF1.
OS   TT virus.
OC   Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX   NCBI_TaxID=68887;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21488921; PubMed=11601907;
RA   Okamoto H., Nishizawa T., Takahashi M., Asabe S., Tsuda F.,
RA   Yoshikawa A.;
RT   "Heterogeneous distribution of TT virus of distinct genotypes in
RT   multiple tissues from infected humans.";
RL   Virology 288:358-368(2001).
DR   EMBL; AB060592; BAB69900.1; -.
DR   InterPro; IPR004219; TTvirus_Unk.
DR   Pfam; PF02956; TT_ORF1; 1.
SQ   SEQUENCE   750 AA;  89223 MW;  616EC86DC3469091 CRC64;

  Query Match             55.6%;  Score 55;  DB 12;  Length 750;
  Best Local Similarity   44.4%;  Pred. No. 38;
  Matches    8;  Conservative    1;  Mismatches    3;  Indels    6;  Gaps    1;

Qy          2 LPWKWPW------WPWRR 13
              : | | |      |||||
Db          1 MAWTWWWQRRRRRWPWRR 18


RESULT 9
Q990M4
ID   Q990M4      PRELIMINARY;      PRT;  1173 AA.
AC   Q990M4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Spike glycoprotein.
GN   S.
OS   Human coronavirus (strain 229E).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC   Coronaviridae; Coronavirus.
OX   NCBI_TaxID=11137;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=229E;
RA   Bonavia A., Holmes K.V.;
RT   "Viral and cellular changes in a human cell line persistently infected
RT   with human coronavirus HCoV-229E.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF344186; AAK32188.1; -.
DR   InterPro; IPR002551; Corona_S1.
DR   InterPro; IPR002552; Corona_S2.
DR   Pfam; PF01600; Corona_S1; 1.
DR   Pfam; PF01601; Corona_S2; 1.
SQ   SEQUENCE   1173 AA;  128669 MW;  ABC6E0A75E8BD8A4 CRC64;

  Query Match             54.5%;  Score 54;  DB 12;  Length 1173;
```

```
  Best Local Similarity   85.7%;  Pred. No. 74;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 KWPWWPW 11
              ||||| |
Db       1113 KWPWWVW 1119


RESULT 10
Q990M1
ID   Q990M1      PRELIMINARY;      PRT;  1173 AA.
AC   Q990M1;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Spike glycoprotein.
GN   S.
OS   Human coronavirus (strain 229E).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC   Coronaviridae; Coronavirus.
OX   NCBI_TaxID=11137;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=229E;
RA   Bonavia A., Holmes K.V.;
RT   "Viral and cellular changes in a human cell line persistently infected
RT   with human coronavirus HCoV-229E.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF344189; AAK32191.1; -.
DR   InterPro; IPR002551; Corona_S1.
DR   InterPro; IPR002552; Corona_S2.
DR   Pfam; PF01600; Corona_S1; 1.
DR   Pfam; PF01601; Corona_S2; 1.
SQ   SEQUENCE   1173 AA;  128760 MW;  B73A165A6270152A CRC64;

  Query Match             54.5%;  Score 54;  DB 12;  Length 1173;
  Best Local Similarity   85.7%;  Pred. No. 74;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 KWPWWPW 11
              ||||| |
Db       1113 KWPWWVW 1119


RESULT 11
Q990M3
ID   Q990M3      PRELIMINARY;      PRT;  1173 AA.
AC   Q990M3;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Spike glycoprotein.
GN   S.
OS   Human coronavirus (strain 229E).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC   Coronaviridae; Coronavirus.
OX   NCBI_TaxID=11137;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=229E;
RA   Bonavia A., Holmes K.V.;
RT   "Viral and cellular changes in a human cell line persistently infected
RT   with human coronavirus HCoV-229E.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF344187; AAK32189.1; -.
DR   InterPro; IPR002551; Corona_S1.
DR   InterPro; IPR002552; Corona_S2.
DR   Pfam; PF01600; Corona_S1; 1.
DR   Pfam; PF01601; Corona_S2; 1.
SQ   SEQUENCE   1173 AA;  128683 MW;  9E2368160082A81A CRC64;

  Query Match             54.5%;  Score 54;  DB 12;  Length 1173;
  Best Local Similarity   85.7%;  Pred. No. 74;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 KWPWWPW 11
              ||||| |
Db       1113 KWPWWVW 1119


RESULT 12
Q990M2
ID   Q990M2      PRELIMINARY;      PRT;  1173 AA.
AC   Q990M2;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Spike glycoprotein.
GN   S.
OS   Human coronavirus (strain 229E).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC   Coronaviridae; Coronavirus.
OX   NCBI_TaxID=11137;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=229E;
RA   Bonavia A., Holmes K.V.;
RT   "Viral and cellular changes in a human cell line persistently infected
RT   with human coronavirus HCoV-229E.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF344188; AAK32190.1; -.
DR   InterPro; IPR002551; Corona_S1.
DR   InterPro; IPR002552; Corona_S2.
DR   Pfam; PF01600; Corona_S1; 1.
DR   Pfam; PF01601; Corona_S2; 1.
SQ   SEQUENCE   1173 AA;  128653 MW;  8B658FCBBD1842DA CRC64;

  Query Match             54.5%;  Score 54;  DB 12;  Length 1173;
  Best Local Similarity   85.7%;  Pred. No. 74;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 KWPWWPW 11
              ||||| |
Db       1113 KWPWWVW 1119


RESULT 13
Q84712
ID   Q84712      PRELIMINARY;      PRT;  1383 AA.
AC   Q84712;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Spike protein.
OS   Porcine epidemic diarrhea virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC   Coronaviridae; Coronavirus.
OX   NCBI_TaxID=28295;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Br1/87;
RX   MEDLINE=94231173; PubMed=8176382;
RA   Duarte M., Laude H.;
RT   "Sequence of the spike protein of the porcine epidemic diarrhoea
RT   virus.";
RL   J. Gen. Virol. 75:1195-1200(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Br1/87;
RX   MEDLINE=93389433; PubMed=8397280;
RA   Bridgen A., Duarte M., Tobler K., Laude H., Ackermann M.;
RT   "Sequence determination of the nucleocapsid protein gene of the
RT   porcine epidemic diarrhoea virus confirms that this virus is a
RT   coronavirus related to human coronavirus 229E and porcine
RT   transmissible gastroenteritis virus.";
RL   J. Gen. Virol. 74:1795-1804(1993).
```

```
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Br1/87;
RX   MEDLINE=94120721; PubMed=8291230;
RA   Duarte M., Tobler K., Bridgen A., Rasschaert D., Ackermann M.,
RA   Laude H.;
RT   "Sequence analysis of the porcine epidemic diarrhea virus genome
RT   between the nucleocapsid and spike protein genes reveals a polymo.";
RL   Virology 198:466-476(1994).
DR   EMBL; Z25483; CAA80971.1; -.
DR   InterPro; IPR002551; Corona_S1.
DR   InterPro; IPR002552; Corona_S2.
DR   Pfam; PF01600; Corona_S1; 1.
DR   Pfam; PF01601; Corona_S2; 1.
FT   CONFLICT    422    422       Y -> N (IN REF. 1).
SQ   SEQUENCE   1383 AA;  151405 MW;  741C84D5DD3BDC4D CRC64;

  Query Match             54.5%;  Score 54;  DB 12;  Length 1383;
  Best Local Similarity   85.7%;  Pred. No. 86;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 KWPWWPW 11
              ||||| |
Db       1322 KWPWWVW 1328


RESULT 14
Q91AV1
ID   Q91AV1      PRELIMINARY;      PRT;  1383 AA.
AC   Q91AV1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Spike protein.
OS   Porcine epidemic diarrhea virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC   Coronaviridae; Coronavirus.
OX   NCBI_TaxID=28295;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV777;
RX   MEDLINE=93389433; PubMed=8397280;
RA   Bridgen A., Duarte M., Tobler K., Laude H., Ackermann M.;
RT   "Sequence determination of the nucleocapsid protein gene of the
RT   porcine epidemic diarrhea virus confirms that this virus is a
RT   coronavirus related to human coronavirus 229E and porcine
RT   transmissible gastroenteritis virus.";
RL   J. Gen. Virol. 74:1795-1804(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV777;
RX   MEDLINE=94120721; PubMed=8291230;
RA   Duarte M., Tobler K., Bridgen A., Rasschaert D., Ackermann M.,
RA   Laude H.;
RT   "Sequence analysis of the porcine epidemic diarrhea virus genome
RT   between the nucleocapsid and spike protein genes reveals a polymorphic
RT   ORF.";
RL   Virology 198:466-476(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV777;
RX   MEDLINE=96112302; PubMed=8830538;
RA   Tobler K., Ackermann M.;
RT   "PEDV leader sequence and junction sites.";
RL   Adv. Exp. Med. Biol. 380:541-542(1995).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV777;
RX   MEDLINE=98455678; PubMed=9782358;
RA   Bridgen A., Kocherhans R., Tobler K., Carvajal A., Ackermann M.;
RT   "Further analysis of the genome of porcine epidemic diarrhea virus.";
RL   Adv. Exp. Med. Biol. 440:781-786(1998).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV777;
RA   Kocherhans R., Bridgen A., Ackermann M., Tobler K.;
RT   "The complete genome sequence of porcine epidemic diarrhea
RT   coronavirus.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF353511; AAK38656.1; -.
DR   InterPro; IPR002551; Corona_S1.
DR   InterPro; IPR002552; Corona_S2.
DR   Pfam; PF01600; Corona_S1; 1.
DR   Pfam; PF01601; Corona_S2; 1.
SQ   SEQUENCE   1383 AA;  151352 MW;  022E5E5E5435876D CRC64;

  Query Match             54.5%;  Score 54;  DB 12;  Length 1383;
  Best Local Similarity   85.7%;  Pred. No. 86;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 KWPWWPW 11
              ||||| |
Db       1322 KWPWWVW 1328


RESULT 15
Q8B482
ID   Q8B482      PRELIMINARY;      PRT;  1383 AA.
AC   Q8B482;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Spike protein.
OS   Porcine epidemic diarrhea virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC   Coronaviridae; Coronavirus.
OX   NCBI_TaxID=28295;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Chinju99;
RA   Yeo S.-G., Krell P., Nagy E.;
RT   "Cloning and nucleotide sequence analysis of spike gene of porcine
RT   epidemic diarrhea virus detected in Korea.";
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY167585; AAN86621.1; -.
SQ   SEQUENCE   1383 AA;  151582 MW;  B5BA4D7EE5371A54 CRC64;

  Query Match             54.5%;  Score 54;  DB 12;  Length 1383;
  Best Local Similarity   85.7%;  Pred. No. 86;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 KWPWWPW 11
              ||||| |
Db       1322 KWPWWVW 1328


Search completed: October  2, 2003, 10:03:57
Job time : 107 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 2, 2003, 10:00:07 ; Search time 25 Seconds
(without alignments)
22.002 Million cell updates/sec

Title: US-09-444-281-85
Perfect score: 99
Sequence: 1 ILPWKWPWWPWRR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	99	100.0	13	1	US-07-715-271-1	Sequence 1, Appli
2	99	100.0	13	1	US-08-197-205-1	Sequence 1, Appli
3	99	100.0	13	1	US-08-197-205-5	Sequence 5, Appli
4	99	100.0	13	3	US-08-915-314-29	Sequence 29, Appl
5	99	100.0	13	3	US-09-230-180-30	Sequence 30, Appl
6	99	100.0	13	3	US-08-702-054B-16	Sequence 16, Appl
7	99	100.0	13	4	US-09-076-227-1	Sequence 1, Appli
8	99	100.0	13	4	US-09-099-631A-1	Sequence 1, Appli
9	99	100.0	13	4	US-09-099-631A-3	Sequence 3, Appli
10	99	100.0	13	4	US-09-318-195A-2	Sequence 2, Appli
11	99	100.0	13	4	US-09-030-619-98	Sequence 98, Appl
12	99	100.0	13	4	US-09-030-619-204	Sequence 204, App
13	99	100.0	13	4	US-09-416-481A-1	Sequence 1, Appli
14	99	100.0	13	4	US-09-667-486-29	Sequence 29, Appl
15	99	100.0	14	4	US-09-076-227-24	Sequence 24, Appl
16	99	100.0	14	4	US-09-076-227-36	Sequence 36, Appl
17	99	100.0	14	4	US-09-416-481A-24	Sequence 24, Appl
18	99	100.0	14	4	US-09-416-481A-36	Sequence 36, Appl
19	99	100.0	15	3	US-08-702-054B-7	Sequence 7, Appli
20	99	100.0	16	4	US-09-076-227-37	Sequence 37, Appl
21	99	100.0	16	4	US-09-416-481A-37	Sequence 37, Appl
22	99	100.0	17	4	US-09-076-227-35	Sequence 35, Appl
23	99	100.0	17	4	US-09-416-481A-35	Sequence 35, Appl
24	99	100.0	63	4	US-09-099-631A-12	Sequence 12, Appl
25	99	100.0	63	4	US-09-416-481A-39	Sequence 39, Appl
26	96	97.0	13	1	US-08-197-205-7	Sequence 7, Appli
27	96	97.0	15	3	US-08-702-054B-6	Sequence 6, Appli
28	95	96.0	12	4	US-09-076-227-2	Sequence 2, Appli
29	95	96.0	12	4	US-09-099-631A-5	Sequence 5, Appli
30	95	96.0	12	4	US-09-416-481A-2	Sequence 2, Appli
31	94	94.9	12	1	US-08-197-205-6	Sequence 6, Appli
32	94	94.9	12	4	US-09-076-227-25	Sequence 25, Appl
33	94	94.9	12	4	US-09-318-195A-6	Sequence 6, Appli
34	94	94.9	12	4	US-09-416-481A-25	Sequence 25, Appl
35	94	94.9	13	1	US-08-197-205-2	Sequence 2, Appli
36	94	94.9	13	4	US-09-318-195A-1	Sequence 1, Appli
37	94	94.9	13	4	US-09-318-195A-3	Sequence 3, Appli
38	94	94.9	13	4	US-09-318-195A-7	Sequence 7, Appli
39	93	93.9	13	1	US-08-197-205-3	Sequence 3, Appli
40	93	93.9	15	3	US-08-702-054B-8	Sequence 8, Appli
41	91	91.9	11	3	US-08-702-054B-9	Sequence 9, Appli
42	91	91.9	11	4	US-09-076-227-3	Sequence 3, Appli
43	91	91.9	11	4	US-09-099-631A-6	Sequence 6, Appli
44	91	91.9	11	4	US-09-416-481A-3	Sequence 3, Appli
45	91	91.9	12	1	US-08-197-205-4	Sequence 4, Appli

ALIGNMENTS

```
RESULT 1
US-07-715-271-1
; Sequence 1, Application US/07715271
; Patent No. 5324716
;  GENERAL INFORMATION:
;    APPLICANT:  Selsted, Michael E.
;    APPLICANT:  Cullor, James S.
;    TITLE OF INVENTION:  BROAD SPECTRUM ANTIMICROBIAL COMPOUNDS
;    TITLE OF INVENTION:  AND METHODS OF USE
;    NUMBER OF SEQUENCES:  1
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PRETTY, SCHROEDER, BRUEGGEMANN & CLARK
;      STREET:  444 So. Flower Street, Suite 2000
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/715,271
;      FILING DATE:  19910614
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P31 8963
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-535-9001
;      TELEFAX:  619-535-8949
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-07-715-271-1

  Query Match            100.0%;  Score 99;  DB 1;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 1.1e-07;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPWRR 13
              |||||||||||||
Db          1 ILPWKWPWWPWRR 13
```

```
RESULT 2
US-08-197-205-1
; Sequence 1, Application US/08197205
; Patent No. 5547939
;  GENERAL INFORMATION:
;    APPLICANT:  Selsted, Michael E.
;    TITLE OF INVENTION:  Broad Spectrum Antimicrobial Compounds
;    TITLE OF INVENTION:  and Methods of Use
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/197,205
;      FILING DATE:  16-FEB-1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UC 9881
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-197-205-1

  Query Match             100.0%;  Score 99;  DB 1;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-07;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPWRR 13
              |||||||||||||
Db          1 ILPWKWPWWPWRR 13


RESULT 3
US-08-197-205-5
; Sequence 5, Application US/08197205
; Patent No. 5547939
;  GENERAL INFORMATION:
;    APPLICANT:  Selsted, Michael E.
;    TITLE OF INVENTION:  Broad Spectrum Antimicrobial Compounds
;    TITLE OF INVENTION:  and Methods of Use
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/197,205
;      FILING DATE:  16-FEB-1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UC 9881
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-197-205-5

  Query Match             100.0%;  Score 99;  DB 1;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-07;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPWRR 13
              |||||||||||||
Db          1 ILPWKWPWWPWRR 13


RESULT 4
US-08-915-314-29
; Sequence 29, Application US/08915314
; Patent No. 6180604
;  GENERAL INFORMATION:
;    APPLICANT:  Fraser, Janet R.
;    APPLICANT:  West, Michael H.P.
;    APPLICANT:  Krieger, Timothy J.
;    APPLICANT:  Taylor, Robert
;    APPLICANT:  Erfle, Douglas
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR TREATING
;    TITLE OF INVENTION:  INFECTIONS USING ANALOGUES OF INDOLICIDIN
;    NUMBER OF SEQUENCES:  90
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  USA
;      ZIP:  98104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/915,314
;      FILING DATE:  20-AUG-1997
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6180604tenburg Ph.D., Carol
;      REGISTRATION NUMBER:  39,317
;      REFERENCE/DOCKET NUMBER:  660081.405
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-915-314-29

  Query Match             100.0%;  Score 99;  DB 3;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-07;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPWRR 13
              |||||||||||||
```

```
Db          1 ILPWKWPWWPWRR 13


RESULT 5
US-09-230-180-30
; Sequence 30, Application US/09230180
; Patent No. 6183992
; GENERAL INFORMATION:
;  APPLICANT:  Kim, Sun-Chang
;  APPLICANT:  Lee, Jae Hyun
;  APPLICANT:  Kang, Min Hyung
;  APPLICANT:  Kim, Jeong Hyun
;  APPLICANT:  Hong, Seung-Suh
;  APPLICANT:  Lee, Hyun-Soo
;  APPLICANT:  Samyang Genex Corporation
;  APPLICANT:  Korea Advanced Institute of Science and Technology
;  TITLE OF INVENTION:  METHOD FOR MASS PRODUCTION OF
;  TITLE OF INVENTION:  ANTIMICROBIAL PEPTIDE
;  FILE REFERENCE:  6181/0F135
;  CURRENT APPLICATION NUMBER: US/09/230,180
;  CURRENT FILING DATE:  1999-03-10
;  PRIOR APPLICATION NUMBER:  PCT/KR98/00132
;  PRIOR FILING DATE:  1998-05-28
;  PRIOR APPLICATION NUMBER:  KR 13372/1998
;  PRIOR FILING DATE:  1998-04-09
;  PRIOR APPLICATION NUMBER:  KR 21312/1997
;  PRIOR FILING DATE:  1997-05-28
;  NUMBER OF SEQ ID NOS:  36
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 30
;   LENGTH: 13
;   TYPE: PRT
;   ORGANISM: Bos taurus
US-09-230-180-30

  Query Match             100.0%;  Score 99;  DB 3;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-07;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPWRR 13
              |||||||||||||
Db          1 ILPWKWPWWPWRR 13


RESULT 6
US-08-702-054B-16
; Sequence 16, Application US/08702054B
; Patent No. 6191254
;  GENERAL INFORMATION:
;    APPLICANT:  Falls, Timothy J.
;    APPLICANT:  Hancock, Robert E. W.
;    APPLICANT:  Gough, Monisha
;    TITLE OF INVENTION:  ANTIMICROBIAL CATIONIC PEPTIDES
;    TITLE OF INVENTION:  AND METHODS OF SCREENING FOR THE SAME
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows 95
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/702,054B
;      FILING DATE:  23-AUG-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/002,687
;      FILING DATE:  23-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Haile, Lisa A.
;      REGISTRATION NUMBER:  38,347
;      REFERENCE/DOCKET NUMBER:  07420/013001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619/678-5070
;      TELEFAX:  619/678-5099
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-702-054B-16

  Query Match             100.0%;  Score 99;  DB 3;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-07;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPWRR 13
              |||||||||||||
Db          1 ILPWKWPWWPWRR 13


RESULT 7
US-09-076-227-1
; Sequence 1, Application US/09076227
; Patent No. 6303575
; GENERAL INFORMATION:
;  APPLICANT: Selsted, Michael E.
;  TITLE OF INVENTION: Indolicidin Analogs and Methods of Using Same
;  FILE REFERENCE: P-UC 3049
;  CURRENT APPLICATION NUMBER: US/09/076,227
;  CURRENT FILING DATE: 1998-05-12
;  NUMBER OF SEQ ID NOS: 37
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 13
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: construct
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (13)
;   OTHER INFORMATION: AMIDATION
US-09-076-227-1

  Query Match             100.0%;  Score 99;  DB 4;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-07;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPWRR 13
              |||||||||||||
Db          1 ILPWKWPWWPWRR 13


RESULT 8
US-09-099-631A-1
; Sequence 1, Application US/09099631A
; Patent No. 6444645
; GENERAL INFORMATION:
;  APPLICANT:  Selsted, Michael E.
;  APPLICANT:  Osapay, Klara
;  TITLE OF INVENTION:  Crosslink-Stabilized Indolicidin Analogs
;  FILE REFERENCE:  P-UC 3050
;  CURRENT APPLICATION NUMBER: US/09/099,631A
;  CURRENT FILING DATE:  1998-06-18
;  NUMBER OF SEQ ID NOS:  13
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 13
```

```
;   TYPE: PRT
;   ORGANISM: Bos taurus
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (13)
;   OTHER INFORMATION: AMIDATION
US-09-099-631A-1

  Query Match             100.0%;  Score 99;  DB 4;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-07;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPWRR 13
              |||||||||||||
Db          1 ILPWKWPWWPWRR 13


RESULT 9
US-09-099-631A-3
; Sequence 3, Application US/09099631A
; Patent No. 6444645
; GENERAL INFORMATION:
;  APPLICANT:  Selsted, Michael E.
;  APPLICANT:  Osapay, Klara
;  TITLE OF INVENTION:  Crosslink-Stabilized Indolicidin Analogs
;  FILE REFERENCE:  P-UC 3050
;  CURRENT APPLICATION NUMBER: US/09/099,631A
;  CURRENT FILING DATE:  1998-06-18
;  NUMBER OF SEQ ID NOS:  13
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 13
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (13)
;   OTHER INFORMATION: AMIDATION
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: Construct
US-09-099-631A-3

  Query Match             100.0%;  Score 99;  DB 4;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-07;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPWRR 13
              |||||||||||||
Db          1 ILPWKWPWWPWRR 13


RESULT 10
US-09-318-195A-2
; Sequence 2, Application US/09318195A
; Patent No. 6482799
; GENERAL INFORMATION:
;  APPLICANT: Tuse, Daniel
;  APPLICANT:  Mortelmans, Kristien
;  APPLICANT:  Hokama, Leslie A.
;  APPLICANT:  Chapoy, Larry L.
;  APPLICANT:  Quinn, Michael H.
;  APPLICANT:  Large Scale Biology Corporation
;  TITLE OF INVENTION: Self-Preserving Multipurpose Ophthalmic Solutions
;  TITLE OF INVENTION:  Incorporating a Polypeptide Antimicrobial
;  FILE REFERENCE: 017942-001400US
;  CURRENT APPLICATION NUMBER: US/09/318,195A
;  CURRENT FILING DATE:  1999-05-25
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 13
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:indolicidin
;   OTHER INFORMATION: analog Indol-12-R13-R-OH
US-09-318-195A-2

  Query Match             100.0%;  Score 99;  DB 4;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-07;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPWRR 13
              |||||||||||||
Db          1 ILPWKWPWWPWRR 13


RESULT 11
US-09-030-619-98
; Sequence 98, Application US/09030619B
; Patent No. 6503881
; GENERAL INFORMATION:
;  APPLICANT:  Krieger, Timothy J.
;  APPLICANT:  Taylor, Robert
;  APPLICANT:  Erfle, Douglas
;  APPLICANT:  Fraser, Janet R.
;  APPLICANT:  West, Michael H.P.
;  APPLICANT:  McNicol, Patricia J.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR TREATING
;  TITLE OF INVENTION:  INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
;  TITLE OF INVENTION:  WITH ANTIBIOTICS
;  FILE REFERENCE:  660081.406
;  CURRENT APPLICATION NUMBER: US/09/030,619B
;  CURRENT FILING DATE:  1998-02-25
;  NUMBER OF SEQ ID NOS:  232
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 98
;   LENGTH: 13
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-98

  Query Match             100.0%;  Score 99;  DB 4;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-07;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPWRR 13
              |||||||||||||
Db          1 ILPWKWPWWPWRR 13


RESULT 12
US-09-030-619-204
; Sequence 204, Application US/09030619B
; Patent No. 6503881
; GENERAL INFORMATION:
;  APPLICANT:  Krieger, Timothy J.
;  APPLICANT:  Taylor, Robert
;  APPLICANT:  Erfle, Douglas
;  APPLICANT:  Fraser, Janet R.
;  APPLICANT:  West, Michael H.P.
;  APPLICANT:  McNicol, Patricia J.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR TREATING
;  TITLE OF INVENTION:  INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
;  TITLE OF INVENTION:  WITH ANTIBIOTICS
;  FILE REFERENCE:  660081.406
;  CURRENT APPLICATION NUMBER: US/09/030,619B
;  CURRENT FILING DATE:  1998-02-25
;  NUMBER OF SEQ ID NOS:  232
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 204
;   LENGTH: 13
;   TYPE: PRT
;   ORGANISM: Bos taurus
```

US-09-030-619-204

Query Match 100.0%; Score 99; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKWPWWPWRR 13
||||||||||||||
Db 1 ILPWKWPWWPWRR 13

RESULT 13
US-09-416-481A-1
; Sequence 1, Application US/09416481A
; Patent No. 6524585
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; TITLE OF INVENTION: Indolicidin Analogs and Methods of Using Same
; FILE REFERENCE: P-UC 3794
; CURRENT APPLICATION NUMBER: US/09/416,481A
; CURRENT FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: US 09/076,227
; PRIOR FILING DATE: 1998-05-12
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: construct
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (13)
; OTHER INFORMATION: AMIDATION
US-09-416-481A-1

Query Match 100.0%; Score 99; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKWPWWPWRR 13
||||||||||||||
Db 1 ILPWKWPWWPWRR 13

RESULT 14
US-09-667-486-29
; Sequence 29, Application US/09667486
; Patent No. 6538106
; GENERAL INFORMATION:
; APPLICANT: Fraser, Janet R.
; West, Michael H.P.
; Krieger, Timothy J.
; Taylor, Robert
; Erfle, Douglas
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; INFECTIONS USING ANALOGUES OF INDOLICIDIN
; NUMBER OF SEQUENCES: 90
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/667,486
; FILING DATE: 22-Sep-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,314
; FILING DATE: 20-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6538106tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 660081.405
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 29:
US-09-667-486-29

Query Match 100.0%; Score 99; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKWPWWPWRR 13
||||||||||||||
Db 1 ILPWKWPWWPWRR 13

RESULT 15
US-09-076-227-24
; Sequence 24, Application US/09076227
; Patent No. 6303575
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; TITLE OF INVENTION: Indolicidin Analogs and Methods of Using Same
; FILE REFERENCE: P-UC 3049
; CURRENT APPLICATION NUMBER: US/09/076,227
; CURRENT FILING DATE: 1998-05-12
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (14)
; OTHER INFORMATION: Xaa is homoserine (Hse).
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: construct
US-09-076-227-24

Query Match 100.0%; Score 99; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKWPWWPWRR 13
||||||||||||||
Db 1 ILPWKWPWWPWRR 13

Search completed: October 2, 2003, 10:05:20
Job time : 26 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 2, 2003, 10:04:03 ; Search time 65 Seconds
(without alignments)
31.643 Million cell updates/sec

Title: US-09-444-281-85
Perfect score: 99
Sequence: 1 ILPWKWPWWPWRR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	99	100.0	13	9	US-09-030-619-98	Sequence 98, Appl
2	99	100.0	13	9	US-09-030-619-204	Sequence 204, App
3	99	100.0	13	9	US-09-917-340-13	Sequence 13, Appl
4	99	100.0	13	11	US-09-988-792-5	Sequence 5, Appli
5	99	100.0	13	11	US-09-820-053A-42	Sequence 42, Appl
6	99	100.0	13	11	US-09-820-053A-57	Sequence 57, Appl
7	99	100.0	13	12	US-10-229-368-3	Sequence 3, Appli
8	99	100.0	13	12	US-10-229-368-4	Sequence 4, Appli
9	99	100.0	13	12	US-10-225-087-3	Sequence 3, Appli
10	99	100.0	13	12	US-10-225-087-4	Sequence 4, Appli
11	99	100.0	13	15	US-10-109-171-42	Sequence 42, Appl
12	99	100.0	13	15	US-10-109-171-57	Sequence 57, Appl
13	99	100.0	13	16	US-10-252-773-2	Sequence 2, Appli
14	99	100.0	19	9	US-09-909-652-5	Sequence 5, Appli
15	81	81.8	12	9	US-09-030-619-40	Sequence 40, Appl
16	81	81.8	12	12	US-10-229-368-17	Sequence 17, Appl
17	81	81.8	12	12	US-10-225-087-17	Sequence 17, Appl
18	75.5	76.3	15	9	US-09-030-619-39	Sequence 39, Appl
19	75.5	76.3	15	12	US-10-229-368-16	Sequence 16, Appl
20	75.5	76.3	15	12	US-10-225-087-16	Sequence 16, Appl
21	75	75.8	28	9	US-09-030-619-50	Sequence 50, Appl
22	75	75.8	28	9	US-09-030-619-104	Sequence 104, App
23	75	75.8	28	12	US-10-229-368-33	Sequence 33, Appl
24	75	75.8	28	12	US-10-225-087-32	Sequence 32, Appl
25	73.5	74.2	13	9	US-09-030-619-107	Sequence 107, App
26	73.5	74.2	13	12	US-10-229-368-40	Sequence 40, Appl
27	73.5	74.2	13	12	US-10-225-087-37	Sequence 37, Appl
28	73	73.7	12	9	US-09-030-619-43	Sequence 43, Appl
29	73	73.7	12	9	US-09-030-619-67	Sequence 67, Appl
30	73	73.7	12	9	US-09-030-619-73	Sequence 73, Appl
31	73	73.7	12	9	US-09-030-619-112	Sequence 112, App
32	73	73.7	12	12	US-10-229-368-20	Sequence 20, Appl
33	73	73.7	12	12	US-10-229-368-41	Sequence 41, Appl
34	73	73.7	12	12	US-10-229-368-82	Sequence 82, Appl
35	73	73.7	12	12	US-10-229-368-87	Sequence 87, Appl
36	73	73.7	12	12	US-10-225-087-20	Sequence 20, Appl
37	73	73.7	12	12	US-10-225-087-38	Sequence 38, Appl
38	73	73.7	12	12	US-10-225-087-72	Sequence 72, Appl
39	73	73.7	12	12	US-10-225-087-77	Sequence 77, Appl
40	73	73.7	13	9	US-09-030-619-53	Sequence 53, Appl
41	73	73.7	13	9	US-09-030-619-95	Sequence 95, Appl
42	73	73.7	13	9	US-09-030-619-99	Sequence 99, Appl
43	73	73.7	13	9	US-09-030-619-109	Sequence 109, App
44	73	73.7	13	12	US-10-229-368-5	Sequence 5, Appli
45	73	73.7	13	12	US-10-229-368-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-09-030-619-98
; Sequence 98, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfle, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 98
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-98

Query Match 100.0%; Score 99; DB 9; Length 13;
Best Local Similarity 100.0%; Pred. No. 3.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKWPWWPWRR 13
||||||||||||||
Db 1 ILPWKWPWWPWRR 13

RESULT 2
US-09-030-619-204

```
; Sequence 204, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
;  APPLICANT:  Krieger, Timothy J.
;  APPLICANT:  Taylor, Robert
;  APPLICANT:  Erfle, Douglas
;  APPLICANT:  Fraser, Janet R.
;  APPLICANT:  West, Michael H.P.
;  APPLICANT:  McNicol, Patricia J.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR TREATING
;  TITLE OF INVENTION:  INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
;  TITLE OF INVENTION:  WITH ANTIBIOTICS
;  FILE REFERENCE:  660081.406
;  CURRENT APPLICATION NUMBER: US/09/030,619B
;  CURRENT FILING DATE:  1998-02-25
;  NUMBER OF SEQ ID NOS:  232
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 204
;   LENGTH: 13
;   TYPE: PRT
;   ORGANISM: Bos taurus
US-09-030-619-204

  Query Match             100.0%;  Score 99;  DB 9;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPWRR 13
              |||||||||||||
Db          1 ILPWKWPWWPWRR 13


RESULT 3
US-09-917-340-13
; Sequence 13, Application US/09917340
; Patent No. US20020090369A1
; GENERAL INFORMATION:
;  APPLICANT: Murphy, Christopher J.
;  APPLICANT:  McAnulty, Jonathan F.
;  APPLICANT:  Reid, Ted W.
;  TITLE OF INVENTION: Transplant Media
;  FILE REFERENCE: TPLANT-06468
;  CURRENT APPLICATION NUMBER: US/09/917,340
;  CURRENT FILING DATE:  2001-07-29
;  PRIOR APPLICATION NUMBER: 60/221,632
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: 60/249,602
;  PRIOR FILING DATE: 2000-11-17
;  PRIOR APPLICATION NUMBER: 60/290,932
;  PRIOR FILING DATE: 2001-05-15
;  NUMBER OF SEQ ID NOS: 96
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
;   LENGTH: 13
;   TYPE: PRT
;   ORGANISM: Bos taurus
US-09-917-340-13

  Query Match             100.0%;  Score 99;  DB 9;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPWRR 13
              |||||||||||||
Db          1 ILPWKWPWWPWRR 13


RESULT 4
US-09-988-792-5
; Sequence 5, Application US/09988792
; Publication No. US20030032599A1
; GENERAL INFORMATION:
;  APPLICANT: Lipkowski, Andrezej W
;  APPLICANT:  Carr, Daniel B
;  TITLE OF INVENTION: NOVEL ANTIMICROBIAL COMPOUNDS
;  FILE REFERENCE: 18475-025
;  CURRENT APPLICATION NUMBER: US/09/988,792
;  CURRENT FILING DATE:  2001-11-20
;  PRIOR APPLICATION NUMBER: 60/252,369
;  PRIOR FILING DATE: 2000-11-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 13
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-988-792-5

  Query Match             100.0%;  Score 99;  DB 11;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPWRR 13
              |||||||||||||
Db          1 ILPWKWPWWPWRR 13


RESULT 5
US-09-820-053A-42
; Sequence 42, Application US/09820053A
; Publication No. US20030083243A1
; GENERAL INFORMATION:
;  APPLICANT: Owen, Donald R.
;  TITLE OF INVENTION: SHORT BIOACTIVE PEPTIDES
;  FILE REFERENCE: HELX027
;  CURRENT APPLICATION NUMBER: US/09/820,053A
;  CURRENT FILING DATE:  2001-03-28
;  NUMBER OF SEQ ID NOS: 165
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42
;   LENGTH: 13
;   TYPE: PRT
;   ORGANISM: ARTIFICIAL SEQUENCE
;   FEATURE:
;   OTHER INFORMATION: SYNTHETIC SEQUENCE
;   NAME/KEY: MOD_RES
;   LOCATION: (13)
;   OTHER INFORMATION: AMIDATION
US-09-820-053A-42

  Query Match             100.0%;  Score 99;  DB 11;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPWRR 13
              |||||||||||||
Db          1 ILPWKWPWWPWRR 13


RESULT 6
US-09-820-053A-57
; Sequence 57, Application US/09820053A
; Publication No. US20030083243A1
; GENERAL INFORMATION:
;  APPLICANT: Owen, Donald R.
;  TITLE OF INVENTION: SHORT BIOACTIVE PEPTIDES
;  FILE REFERENCE: HELX027
;  CURRENT APPLICATION NUMBER: US/09/820,053A
;  CURRENT FILING DATE:  2001-03-28
;  NUMBER OF SEQ ID NOS: 165
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
;   LENGTH: 13
;   TYPE: PRT
;   ORGANISM: ARTIFICIAL SEQUENCE
;   FEATURE:
```

; OTHER INFORMATION: SYNTHETIC SEQUENCE
US-09-820-053A-57

Query Match 100.0%; Score 99; DB 11; Length 13;
Best Local Similarity 100.0%; Pred. No. 3.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKWPWWPWRR 13
||||||||||||||
Db 1 ILPWKWPWWPWRR 13

RESULT 7
US-10-229-368-3
; Sequence 3, Application US/10229368
; Publication No. US20030148945A1
; GENERAL INFORMATION:
; APPLICANT: McNicol, Patricia J.
; APPLICANT: Pawlak, Sonia K.
; APPLICANT: Rubinchik, Evelina
; APPLICANT: Cameron, Dale
; APPLICANT: Guarna, Maria Marta
; TITLE OF INVENTION: ANTIMICROBIAL AND ANTI-INFLAMMATORY
; TITLE OF INVENTION: PEPTIDES
; FILE REFERENCE: 660081.418
; CURRENT APPLICATION NUMBER: US/10/229,368
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin peptide analogs
US-10-229-368-3

Query Match 100.0%; Score 99; DB 12; Length 13;
Best Local Similarity 100.0%; Pred. No. 3.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKWPWWPWRR 13
||||||||||||||
Db 1 ILPWKWPWWPWRR 13

RESULT 8
US-10-229-368-4
; Sequence 4, Application US/10229368
; Publication No. US20030148945A1
; GENERAL INFORMATION:
; APPLICANT: McNicol, Patricia J.
; APPLICANT: Pawlak, Sonia K.
; APPLICANT: Rubinchik, Evelina
; APPLICANT: Cameron, Dale
; APPLICANT: Guarna, Maria Marta
; TITLE OF INVENTION: ANTIMICROBIAL AND ANTI-INFLAMMATORY
; TITLE OF INVENTION: PEPTIDES
; FILE REFERENCE: 660081.418
; CURRENT APPLICATION NUMBER: US/10/229,368
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin peptide analogs
US-10-229-368-4

Query Match 100.0%; Score 99; DB 12; Length 13;
Best Local Similarity 100.0%; Pred. No. 3.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKWPWWPWRR 13
||||||||||||||
Db 1 ILPWKWPWWPWRR 13

RESULT 9
US-10-225-087-3
; Sequence 3, Application US/10225087
; Publication No. US20030171281A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: McNicol, Patrica J.
; APPLICANT: Fraser, Janet R.
; TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES AND
; TITLE OF INVENTION: FORMULATIONS THEREOF
; FILE REFERENCE: 660081.417
; CURRENT APPLICATION NUMBER: US/10/225,087
; CURRENT FILING DATE: 2003-01-10
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin analog
US-10-225-087-3

Query Match 100.0%; Score 99; DB 12; Length 13;
Best Local Similarity 100.0%; Pred. No. 3.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKWPWWPWRR 13
||||||||||||||
Db 1 ILPWKWPWWPWRR 13

RESULT 10
US-10-225-087-4
; Sequence 4, Application US/10225087
; Publication No. US20030171281A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: McNicol, Patrica J.
; APPLICANT: Fraser, Janet R.
; TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES AND
; TITLE OF INVENTION: FORMULATIONS THEREOF
; FILE REFERENCE: 660081.417
; CURRENT APPLICATION NUMBER: US/10/225,087
; CURRENT FILING DATE: 2003-01-10
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Indolicidin analog
US-10-225-087-4

Query Match 100.0%; Score 99; DB 12; Length 13;
Best Local Similarity 100.0%; Pred. No. 3.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILPWKWPWWPWRR 13
||||||||||||||
Db 1 ILPWKWPWWPWRR 13

RESULT 11
US-10-109-171-42

```
; Sequence 42, Application US/10109171
; Publication No. US20030109452A1
; GENERAL INFORMATION:
;  APPLICANT: Owen, Donald R.
;  TITLE OF INVENTION: SHORT BIOACTIVE PEPTIDES AND METHODS FOR THEIR USE
;  FILE REFERENCE: HELX028
;  CURRENT APPLICATION NUMBER: US/10/109,171
;  CURRENT FILING DATE:  2002-03-28
;  NUMBER OF SEQ ID NOS: 165
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42
;   LENGTH: 13
;   TYPE: PRT
;   ORGANISM: ARTIFICIAL SEQUENCE
;   FEATURE:
;   OTHER INFORMATION: SYNTHETIC SEQUENCE
;   NAME/KEY: MOD_RES
;   LOCATION: (13)
;   OTHER INFORMATION: AMIDATION
US-10-109-171-42

  Query Match             100.0%;  Score 99;  DB 15;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPWRR 13
              |||||||||||||
Db          1 ILPWKWPWWPWRR 13


RESULT 12
US-10-109-171-57
; Sequence 57, Application US/10109171
; Publication No. US20030109452A1
; GENERAL INFORMATION:
;  APPLICANT: Owen, Donald R.
;  TITLE OF INVENTION: SHORT BIOACTIVE PEPTIDES AND METHODS FOR THEIR USE
;  FILE REFERENCE: HELX028
;  CURRENT APPLICATION NUMBER: US/10/109,171
;  CURRENT FILING DATE:  2002-03-28
;  NUMBER OF SEQ ID NOS: 165
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
;   LENGTH: 13
;   TYPE: PRT
;   ORGANISM: ARTIFICIAL SEQUENCE
;   FEATURE:
;   OTHER INFORMATION: SYNTHETIC SEQUENCE
US-10-109-171-57

  Query Match             100.0%;  Score 99;  DB 15;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPWRR 13
              |||||||||||||
Db          1 ILPWKWPWWPWRR 13


RESULT 13
US-10-252-773-2
; Sequence 2, Application US/10252773
; Publication No. US20030131383A1
; GENERAL INFORMATION:
;  APPLICANT: EVERETT, NICHOLAS P.
;  APPLICANT:  LI, QUNIGSHUN
;  APPLICANT:  LAWRENCE, CHRISTOPHER
;  APPLICANT:  DAVIES, MAELOR H.
;  TITLE OF INVENTION: PEPTIDES WITH ENHANCED STABILITY TO PROTEASE
;  TITLE OF INVENTION:  DEGRADATION
;  FILE REFERENCE: INTERLINK 3.0-003
;  CURRENT APPLICATION NUMBER: US/10/252,773
;  CURRENT FILING DATE:  2002-09-23
;  PRIOR APPLICATION NUMBER: 60/106,373
;  PRIOR FILING DATE: 1998-10-30
;  PRIOR APPLICATION NUMBER: 60/106,573
;  PRIOR FILING DATE: 1998-11-02
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 13
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: antimicrobial peptide
US-10-252-773-2

  Query Match             100.0%;  Score 99;  DB 16;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPWRR 13
              |||||||||||||
Db          1 ILPWKWPWWPWRR 13


RESULT 14
US-09-909-652-5
; Sequence 5, Application US/09909652
; Patent No. US20020025537A1
; GENERAL INFORMATION:
;  APPLICANT: Kairos Scientific, Inc.
;  APPLICANT:  Bylina, Edward J.
;  APPLICANT:  Coleman, William J.
;  APPLICANT:  Youvan, Douglas C.
;  TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR GENERATING
;  TITLE OF INVENTION:  AND SCREENING COMPOUNDS THAT AFFECT CELL VIABILITY
;  FILE REFERENCE: 22346-7001
;  CURRENT APPLICATION NUMBER: US/09/909,652
;  CURRENT FILING DATE:  2001-10-15
;  PRIOR APPLICATION NUMBER: US 60/219,179
;  PRIOR FILING DATE: 2000-07-19
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 19
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Ubiquitin indolicidin fusion protein fragment
US-09-909-652-5

  Query Match             100.0%;  Score 99;  DB 9;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 4.8e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILPWKWPWWPWRR 13
              |||||||||||||
Db          7 ILPWKWPWWPWRR 19


RESULT 15
US-09-030-619-40
; Sequence 40, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
;  APPLICANT:  Krieger, Timothy J.
;  APPLICANT:  Taylor, Robert
;  APPLICANT:  Erfle, Douglas
;  APPLICANT:  Fraser, Janet R.
;  APPLICANT:  West, Michael H.P.
;  APPLICANT:  McNicol, Patricia J.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR TREATING
;  TITLE OF INVENTION:  INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
;  TITLE OF INVENTION:  WITH ANTIBIOTICS
```

```
;  FILE REFERENCE:  660081.406
;  CURRENT APPLICATION NUMBER: US/09/030,619B
;  CURRENT FILING DATE:  1998-02-25
;  NUMBER OF SEQ ID NOS:  232
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 40
;   LENGTH: 12
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Indolicidin Analogue
US-09-030-619-40

  Query Match             81.8%;  Score 81;  DB 9;  Length 12;
  Best Local Similarity   90.0%;  Pred. No. 0.0033;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 WKWPWWPWRR 13
              ¦:||||||||
Db          2 WRWPWWPWRR 11


Search completed: October  2, 2003, 10:12:53
Job time : 66 secs
```